PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/56, 15/63, 1/21, 9/24, 15/11, A61K 38/47

A1

(11) International Publication Number: WO 99/11798

(43) International Publication Date: 11 March 1999 (11.03.99)

(21) International Application Number: PCT/US98/17954

(22) International Filing Date: 31 August 1998 (31.08.98)

(30) Priority Data:

08/922,170	2 September 1997 (02.09.97)	US
09/109,386	2 July 1998 (02.07.98)	US

(71) Applicants *(for all designated States except US)*: INSIGHT STRATEGY & MARKETING LTD. [IL/IL]; Kiryat Weizmann Science Park, P.O. Box 2128, 76121 Rehovot (IL). HADASIT MEDICAL RESEARCH SERVICES & DEVELOPMENT LTD. [IL/IL]; Kiryat Hadassah, P.O. Box 12000, 91120 Jerusalem (IL).

(71) Applicant *(for TJ only)*: FRIEDMAN, Mark, M. [US/IL]; Alharizi 1, 43406 Raanana (IL).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: PECKER, Iris [IL/IL]; Wolfson Street 42, 75203 Rishon Le Zion (IL). VLODAVSKY, Israel [IL/IL]; Arbel Street 34, 90805 Mevaseret Zion (IL). FEINSTEIN, Elena [IL/IL]; Hahagana Street 2/29, 76214 Rehovot (IL).

(74) Common Representative: FRIEDMAN, Mark, M.; c/o Castorina, Anthony, Suite 207, 2001 Jefferson Davis Highway, Arlington, VA 22202 (IL).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY AND EXPRESSION OF SAME IN TRANSDUCED CELLS

(57) Abstract

A polynucleotide (*hpa*) encoding a polypeptide having heparanase activity, vectors including same, transduced cells expressing heparanase and a recombinant protein having heparanase activity.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY AND EXPRESSION OF SAME IN TRANSDUCED CELLS

FIELD AND BACKGROUND OF THE INVENTION

The present invention relates to a polynucleotide, referred to hereinbelow as *hpa*, encoding a polypeptide having heparanase activity, vectors including same and transduced cells expressing heparanase. The invention further relates to a recombinant protein having heparanase activity.

Heparan sulfate proteoglycans: Heparan sulfate proteoglycans (HSPG) are ubiquitous macromolecules associated with the cell surface and extra cellular matrix (ECM) of a wide range of cells of vertebrate and invertebrate tissues (1-4). The basic HSPG structure includes a protein core to which several linear heparan sulfate chains are covalently attached. These polysaccharide chains are typically composed of repeating hexuronic and D-glucosamine disaccharide units that are substituted to a varying extent with N- and O-linked sulfate moieties and N-linked acetyl groups (1-4). Studies on the involvement of ECM molecules in cell attachment, growth and differentiation revealed a central role of HSPG in embryonic morphogenesis, angiogenesis, neurite outgrowth and tissue repair (1-5). HSPG are prominent components of blood vessels (3). In large blood vessels they are concentrated mostly in the intima and inner media, whereas in capillaries they are found mainly in the subendothelial basement membrane where they support proliferating and migrating endothelial cells and stabilize the structure of the capillary wall. The ability of HSPG to interact with ECM macromolecules such as collagen, laminin and fibronectin, and with different attachment sites on plasma membranes suggests a key role for this proteoglycan in the self-assembly and insolubility of ECM components, as well as in cell adhesion and locomotion. Cleavage of the heparan sulfate (HS) chains may therefore result in degradation of the subendothelial ECM and hence may play a decisive role in extravasation of blood-borne cells. HS catabolism is observed in inflammation, wound repair, diabetes, and cancer metastasis, suggesting that enzymes which degrade HS play important roles in pathologic processes. Heparanase activity has been described in activated immune system cells and highly metastatic cancer cells (6-8), but research has been handicapped by the lack of biologic tools to explore potential causative roles of heparanase in disease conditions.

Involvement of Heparanase in Tumor Cell Invasion and Metastasis: Circulating tumor cells arrested in the capillary beds of different organs must invade the endothelial cell lining and degrade its underlying basement membrane (BM) in order to invade into the extravascular tissue(s) where they establish

metastasis (9, 10). Metastatic tumor cells often attach at or near the intercellular junctions between adjacent endothelial cells. Such attachment of the metastatic cells is followed by rupture of the junctions, retraction of the endothelial cell borders and migration through the breach in the endothelium toward the exposed underlying BM (9). Once located between endothelial cells and the BM, the invading cells must degrade the subendothelial glycoproteins and proteoglycans of the BM in order to migrate out of the vascular compartment. Several cellular enzymes (e.g., collagenase IV, plasminogen activator, cathepsin B, elastase, etc.) are thought to be involved in degradation of BM (10). Among these enzymes is an endo-β-D-glucuronidase (heparanase) that cleaves HS at specific intrachain sites (6, 8, 11). Expression of a HS degrading heparanase was found to correlate with the metastatic potential of mouse lymphoma (11), fibrosarcoma and melanoma (8) cells. Moreover, elevated levels of heparanase were detected in sera from metastatic tumor bearing animals and melanoma patients (8) and in tumor biopsies of cancer patients (12).

The control of cell proliferation and tumor progression by the local microenvironment, focusing on the interaction of cells with the extracellular matrix (ECM) produced by cultured corneal and vascular endothelial cells, was investigated previously by the present inventors. This cultured ECM closely resembles the subendothelium *in vivo* in its morphological appearance and molecular composition. It contains collagens (mostly type III and IV, with smaller amounts of types I and V), proteoglycans (mostly heparan sulfate- and dermatan sulfate- proteoglycans, with smaller amounts of chondroitin sulfate proteoglycans), laminin, fibronectin, entactin and elastin (13, 14). The ability of cells to degrade HS in the cultured ECM was studied by allowing cells to interact with a metabolically sulfate labeled ECM, followed by gel filtration (Sepharose 6B) analysis of degradation products released into the culture medium (11). While intact HSPG are eluted next to the void volume of the column ($Kav<0.2$, $Mr \sim 0.5x10^6$), labeled degradation fragments of HS side chains are eluted more toward the V_t of the column ($0.5<kav<0.8$, $Mr = 5-7x10^3$) (11).

The heparanase inhibitory effect of various non-anticoagulant species of heparin that might be of potential use in preventing extravasation of blood-borne cells was also investigated by the present inventors. Inhibition of heparanase was best achieved by heparin species containing 16 sugar units or more and having sulfate groups at both the N and O positions. While O-desulfation abolished the heparanase inhibiting effect of heparin, O-sulfated, N-acetylated heparin retained a high inhibitory activity, provided that the N-substituted molecules had a molecular size of about 4,000 daltons or more (7). Treatment of experimental

animals with heparanase inhibitors (e.g., non-anticoagulant species of heparin) markedly reduced (>90%) the incidence of lung metastases induced by B16 melanoma, Lewis lung carcinoma and mammary adenocarcinoma cells (7, 8, 16). Heparin fractions with high and low affinity to anti-thrombin III exhibited a comparable high anti-metastatic activity, indicating that the heparanase inhibiting activity of heparin, rather than its anticoagulant activity, plays a role in the anti-metastatic properties of the polysaccharide (7).

Heparanase activity in the urine of cancer patients: In an attempt to further elucidate the involvement of heparanase in tumor progression and its relevance to human cancer, urine samples for heparanase activity were screened (16a). Heparanase activity was detected in the urine of some, but not all, cancer patients. High levels of heparanase activity were determined in the urine of patients with an aggressive metastatic disease and there was no detectable activity in the urine of healthy donors.

Heparanase activity was also found in the urine of 20% of normal and microalbuminuric insulin dependent diabetes mellitus (IDDM) patients, most likely due to diabetic nephropathy, the most important single disorder leading to renal failure in adults.

Possible involvement of heparanase in tumor angiogenesis: Fibroblast growth factors are a family of structurally related polypeptides characterized by high affinity to heparin (17). They are highly mitogenic for vascular endothelial cells and are among the most potent inducers of neovascularization (17, 18). Basic fibroblast growth factor (bFGF) has been extracted from the subendothelial ECM produced *in vitro* (19) and from basement membranes of the cornea (20), suggesting that ECM may serve as a reservoir for bFGF. Immunohistochemical staining revealed the localization of bFGF in basement membranes of diverse tissues and blood vessels (21). Despite the ubiquitous presence of bFGF in normal tissues, endothelial cell proliferation in these tissues is usually very low, suggesting that bFGF is somehow sequestered from its site of action. Studies on the interaction of bFGF with ECM revealed that bFGF binds to HSPG in the ECM and can be released in an active form by HS degrading enzymes (15, 20, 22). It was demonstrated that heparanase activity expressed by platelets, mast cells, neutrophils, and lymphoma cells is involved in release of active bFGF from ECM and basement membranes (23), suggesting that heparanase activity may not only function in cell migration and invasion, but may also elicit an indirect neovascular response. These results suggest that the ECM HSPG provides a natural storage depot for bFGF and possibly other heparin-binding growth promoting factors (24, 25). Displacement of bFGF from its storage within

basement membranes and ECM may therefore provide a novel mechanism for induction of neovascularization in normal and pathological situations.

Recent studies indicate that heparin and HS are involved in binding of bFGF to high affinity cell surface receptors and in bFGF cell signaling (26, 27). Moreover, the size of HS required for optimal effect was similar to that of HS fragments released by heparanase (28). Similar results were obtained with vascular endothelial cells growth factor (VEGF) (29), suggesting the operation of a dual receptor mechanism involving HS in cell interaction with heparin-binding growth factors. It is therefore proposed that restriction of endothelial cell growth factors in ECM prevents their systemic action on the vascular endothelium, thus maintaining a very low rate of endothelial cells turnover and vessel growth. On the other hand, release of bFGF from storage in ECM as a complex with HS fragment, may elicit localized endothelial cell proliferation and neovascularization in processes such as wound healing, inflammation and tumor development (24, 25).

Expression of heparanase by cells of the immune system: Heparanase activity correlates with the ability of activated cells of the immune system to leave the circulation and elicit both inflammatory and autoimmune responses. Interaction of platelets, granulocytes, T and B lymphocytes, macrophages and mast cells with the subendothelial ECM is associated with degradation of HS by a specific heparanase activity (6). The enzyme is released from intracellular compartments (e.g., lysosomes, specific granules, etc.) in response to various activation signals (e.g., thrombin, calcium ionophore, immune complexes, antigens, mitogens, etc.), suggesting its regulated involvement in inflammation and cellular immunity.

Some of the observations regarding the heparanase enzyme were reviewed in reference No. 6 and are listed hereinbelow:

First, a proteolytic activity (plasminogen activator) and heparanase participate synergistically in sequential degradation of the ECM HSPG by inflammatory leukocytes and malignant cells.

Second, a large proportion of the platelet heparanase exists in a latent form, probably as a complex with chondroitin sulfate. The latent enzyme is activated by tumor cell-derived factor(s) and may then facilitate cell invasion through the vascular endothelium in the process of tumor metastasis.

Third, release of the platelet heparanase from α-granules is induced by a strong stimulant (i.e., thrombin), but not in response to platelet activation on ECM.

Fourth, the neutrophil heparanase is preferentially and readily released in response to a threshold activation and upon incubation of the cells on ECM.

Fifth, contact of neutrophils with ECM inhibited release of noxious enzymes (proteases, lysozyme) and oxygen radicals, but not of enzymes (heparanase, gelatinase) which may enable diapedesis. This protective role of the subendothelial ECM was observed when the cells were stimulated with soluble factors but not with phagocytosable stimulants.

Sixth, intracellular heparanase is secreted within minutes after exposure of T cell lines to specific antigens.

Seventh, mitogens (Con A, LPS) induce synthesis and secretion of heparanase by normal T and B lymphocytes maintained *in vitro*. T lymphocyte heparanase is also induced by immunization with antigen *in vivo*.

Eighth, heparanase activity is expressed by pre-B lymphomas and B-lymphomas, but not by plasmacytomas and resting normal B lymphocytes.

Ninth, heparanase activity is expressed by activated macrophages during incubation with ECM, but there was little or no release of the enzyme into the incubation medium. Similar results were obtained with human myeloid leukemia cells induced to differentiate to mature macrophages.

Tenth, T-cell mediated delayed type hypersensitivity and experimental autoimmunity are suppressed by low doses of heparanase inhibiting non-anticoagulant species of heparin (30).

Eleventh, heparanase activity expressed by platelets, neutrophils and metastatic tumor cells releases active bFGF from ECM and basement membranes. Release of bFGF from storage in ECM may elicit a localized neovascular response in processes such as wound healing, inflammation and tumor development.

Twelfth, among the breakdown products of the ECM generated by heparanase is a tri-sulfated disaccharide that can inhibit T-cell mediated inflammation *in vivo* (31). This inhibition was associated with an inhibitory effect of the disaccharide on the production of biologically active TNFα by activated T cells *in vitro* (31).

Other potential therapeutic applications: Apart from its involvement in tumor cell metastasis, inflammation and autoimmunity, mammalian heparanase may be applied to modulate: bioavailability of heparin-binding growth factors (15); cellular responses to heparin-binding growth factors (e.g., bFGF, VEGF) and cytokines (IL-8) (31a, 29); cell interaction with plasma lipoproteins (32); cellular susceptibility to certain viral and some bacterial and protozoa infections (33, 33a, 33b); and disintegration of amyloid plaques (34). Heparanase may thus

prove useful for conditions such as wound healing, angiogenesis, restenosis, atherosclerosis, inflammation, neurodegenerative diseases and viral infections. Mammalian heparanase can be used to neutralize plasma heparin, as a potential replacement of protamine. Anti-heparanase antibodies may be applied for immunodetection and diagnosis of micrometastases, autoimmune lesions and renal failure in biopsy specimens, plasma samples, and body fluids. Common use in basic research is expected.

The identification of the *hpa* gene encoding for heparanase enzyme will enable the production of a recombinant enzyme in heterologous expression systems. Availability of the recombinant protein will pave the way for solving the protein structure function relationship and will provide a tool for developing new inhibitors.

Viral Infection: The presence of heparan sulfate on cell surfaces have been shown to be the principal requirement for the binding of Herpes Simplex (33) and Dengue (33a) viruses to cells and for subsequent infection of the cells. Removal of the cell surface heparan sulfate by heparanase may therefore abolish virus infection. In fact, treatment of cells with bacterial heparitinase (degrading heparan sulfate) or heparinase (degrading heparan) reduced the binding of two related animal herpes viruses to cells and rendered the cells at least partially resistant to virus infection (33). There are some indications that the cell surface heparan sulfate is also involved in HIV infection (33b).

Neurodegenerative diseases: Heparan sulfate proteoglycans were identified in the prion protein amyloid plaques of Genstmann-Straussler Syndrome, Creutzfeldt-Jakob disease and Scrape (34). Heparanase may disintegrate these amyloid plaques which are also thought to play a role in the pathogenesis of Alzheimer's disease.

Restenosis and Atherosclerosis: Proliferation of arterial smooth muscle cells (SMCs) in response to endothelial injury and accumulation of cholesterol rich lipoproteins are basic events in the pathogenesis of atherosclerosis and restenosis (35). Apart from its involvement in SMC proliferation (i.e., low affinity receptors for heparin-binding growth factors), HS is also involved in lipoprotein binding, retention and uptake (36). It was demonstrated that HSPG and lipoprotein lipase participate in a novel catabolic pathway that may allow substantial cellular and interstitial accumulation of cholesterol rich lipoproteins (32). The latter pathway is expected to be highly atherogenic by promoting accumulation of apoB and apoE rich lipoproteins (i.e. LDL, VLDL, chylomicrons), independent of feed back inhibition by the cellular sterol content. Removal of SMC HS by heparanase is therefore expected to inhibit both SMC

proliferation and lipid accumulation and thus may halt the progression of restenosis and atherosclerosis.

There is thus a widely recognized need for, and it would be highly advantageous to have a polynucleotide encoding a polypeptide having heparanase activity, vectors including same, transduced cells expressing heparanase and a recombinant protein having heparanase activity.

SUMMARY OF THE INVENTION

According to the present invention there is provided a polynucleotide, referred to hereinbelow as *hpa*, *hpa* cDNA or *hpa* gene, encoding a polypeptide having heparanase activity, vectors including same, transduced cells expressing heparanase and a recombinant protein having heparanase activity.

Cloning of the human *hpa* gene which encodes heparanase, and expression of recombinant heparanase by transfected host cells is reported.

A purified preparation of heparanase isolated from human hepatoma cells was subjected to tryptic digestion and microsequencing. The YGPDVGQPR (SEQ ID NO:8) sequence revealed was used to screen EST databases for homology to the corresponding back translated DNA sequence. Two closely related EST sequences were identified and were thereafter found to be identical. Both clones contained an insert of 1020 bp which included an open reading frame of 973 bp followed by a 27 bp of 3' untranslated region and a Poly A tail. Translation start site was not identified.

Cloning of the missing 5' end of *hpa* was performed by PCR amplification of DNA from placenta Marathon RACE cDNA composite using primers selected according to the EST clones sequence and the linkers of the composite. A 900 bp PCR fragment, partially overlapping with the identified 3' encoding EST clones was obtained. The joined cDNA fragment (*hpa*), 1721 bp long (SEQ ID NO:9), contained an open reading frame which encodes a polypeptide of 543 amino acids (SEQ ID NO:10) with a calculated molecular weight of 61,192 daltons.

Cloning an extended 5' sequence was enabled from the human SK-hep1 cell line by PCR amplification using the Marathon RACE. The 5' extended sequence of the SK-hep1 *hpa* cDNA was assembled with the sequence of the *hpa* cDNA isolated from human placenta (SEQ ID NO:9). The assembled sequence contained an open reading frame, SEQ ID NOs: 13 and 15, which encodes, as shown in SEQ ID NOs:14 and 15, a polypeptide of 592 amino acids with a calculated molecular weight of 66,407 daltons.

The ability of the *hpa* gene product to catalyze degradation of heparan sulfate in an *in vitro* assay was examined by expressing the entire open reading

frame of *hpa* in insect cells, using the Baculovirus expression system. Extracts and conditioned media of cells infected with virus containing the *hpa* gene, demonstrated a high level of heparan sulfate degradation activity both towards soluble ECM-derived HSPG and intact ECM. This degradation activity was inhibited by heparin, which is another substrate of heparanase. Cells infected with a similar construct containing no *hpa* gene had no such activity, nor did non-infected cells. The ability of heparanase expressed from the extended 5' clone towards heparin was demonstrated in a mammalian expression system.

The expression pattern of *hpa* RNA in various tissues and cell lines was investigated using RT-PCR. It was found to be expressed only in tissues and cells previously known to have heparanase activity.

A panel of monochromosomal human/CHO and human/mouse somatic cell hybrids was used to localize the human heparanase gene to human chromosome 4. The newly isolated heparanase sequence can be used to identify a chromosome region harboring a human heparanase gene in a chromosome spread.

According to further features in preferred embodiments of the invention described below, there is provided a polynucleotide fragment which includes a polynucleotide sequence encoding a polypeptide having heparanase catalytic activity.

According to still further features in the described preferred embodiments the polynucleotide fragment includes nucleotides 63-1691 of SEQ ID NO:9 or nucleotides 139-1869 of SEQ ID NO:13, which encode the entire human heparanase enzyme.

According to still further features in the described preferred embodiments there is provided a polynucleotide fragment which includes a polynucleotide sequence capable of hybridizing with *hpa* cDNA, especially with nucleotides 1-721 of SEQ ID NO:9.

According to still further features in the described preferred embodiments the polynucleotide sequence which encodes the polypeptide having heparanase activity shares at least 60 % homology, preferably at least 70 % homology, more preferably at least 80 % homology, most preferably at least 90 % homology with SEQ ID NOs:9 or 13.

According to still further features in the described preferred embodiments the polynucleotide fragment according to the present invention includes a portion (fragment) of SEQ ID NOs:9, or 13. For example, such fragments could include nucleotides 63-721 of SEQ ID NO:9 and/or a segment of SEQ ID NO:9 which encodes a polypeptide having the heparanase catalytic activity.

According to still further features in the described preferred embodiments the polypeptide encoded by the polynucleotide fragment includes an amino acid sequence as set forth in SEQ ID NOs:10 or 14 or a functional part thereof.

According to still further features in the described preferred embodiments the polynucleotide sequence encodes a polypeptide having heparanase activity, which shares at least 60 % homology, preferably at least 70 % homology, more preferably at least 80 % homology, most preferably at least 90 % homology with SEQ ID NOs:10 or 14.

According to still further features in the described preferred embodiments the polynucleotide fragment encodes a polypeptide having heparanase activity, which may therefore be allelic, species and/or induced variant of the amino acid sequence set forth in SEQ ID NOs:10 or 14. It is understood that any such variant may also be considered a homolog.

According to still further features in the described preferred embodiments there is provided a single stranded polynucleotide fragment which includes a polynucleotide sequence complementary to at least a portion of a polynucleotide strand encoding a polypeptide having heparanase catalytic activity as described above.

According to still further features in the described preferred embodiments there is provided a vector including a polynucleotide sequence encoding a polypeptide having heparanase catalytic activity.

The vector may be of any suitable type including but not limited to a phage, virus, plasmid, phagemid, cosmid, bacmid or even an artificial chromosome. The polynucleotide sequence encoding a polypeptide having heparanase catalytic activity may include any of the above described polynucleotide fragments.

According to still further features in the described preferred embodiments there is provided a host cell which includes an exogenous polynucleotide fragment including a polynucleotide sequence encoding a polypeptide having heparanase catalytic activity.

The exogenous polynucleotide fragment may be any of the above described fragments. The host cell may be of any type such as prokaryotic cell, eukaryotic cell, a cell line, or a cell as a portion of a multicellular organism (e.g., cells of a transgenic organism).

According to still further features in the described preferred embodiments there is provided a recombinant protein including a polypeptide having heparanase catalytic activity.

According to still further features in the described preferred embodiments there is provided a pharmaceutical composition comprising as an active ingredient a recombinant protein having heparanase catalytic activity.

According to still further features in the described preferred embodiments there is provided a medical equipment comprising a medical device containing, as an active ingredient a recombinant protein having heparanase catalytic activity.

According to still further features in the described preferred embodiments there is provided a heparanase overexpression system comprising a cell overexpressing heparanase catalytic activity.

According to still further features in the described preferred embodiments there is provided a method of identifying a chromosome region harboring a human heparanase gene in a chromosome spread comprising the steps of (a) hybridizing the chromosome spread with a tagged polynucleotyde probe encoding heparanase; (b) washing the chromosome spread, thereby removing excess of non-hybridized probe; and (c) searching for signals associated with said hybridized tagged polynucleotyde probe, wherein detected signals being indicative of a chromosome region harboring a human heparanase gene.

The present invention can be used to develop new drugs to inhibit tumor cell metastasis, inflammation and autoimmunity. The identification of the *hpa* gene encoding for heparanase enzyme enables the production of a recombinant enzyme in heterologous expression systems.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention herein described, by way of example only, with reference to the accompanying drawings, wherein:

FIG. 1 presents nucleotide sequence and deduced amino acid sequence of *hpa* cDNA. A single nucleotide difference at position 799 (A to T) between the EST (Expressed Sequence Tag) and the PCR amplified cDNA (reverse transcribed RNA) and the resulting amino acid substitution (Tyr to Phe) are indicated above and below the substituted unit, respectively. Cysteine residues and the poly adenylation consensus sequence are underlined. The asterisk denotes the stop codon TGA.

FIG. 2 demonstrates degradation of soluble sulfate labeled HSPG substrate by lysates of High Five cells infected with pF*hpa*2 virus. Lysates of High Five cells that were infected with pF*hpa*2 virus (•) or control pF2 virus (□) were incubated (18 h, 37 °C) with sulfate labeled ECM-derived soluble HSPG (peak I). The incubation medium was then subjected to gel filtration on Sepharose 6B.

Low molecular weight HS degradation fragments (peak II) were produced only during incubation with the pF*hpa*2 infected cells, but there was no degradation of the HSPG substrate (⟡) by lysates of pF2 infected cells.

FIGs. 3a-b demonstrate degradation of soluble sulfate labeled HSPG substrate by the culture medium of pF*hpa*2 and pF*hpa*4 infected cells. Culture media of High Five cells infected with pF*hpa*2 (3a) or pF*hpa*4 (3b) viruses (•), or with control viruses (□) were incubated (18 h, 37 °C) with sulfate labeled ECM-derived soluble HSPG (peak I, ⟡). The incubation media were then subjected to gel filtration on Sepharose 6B. Low molecular weight HS degradation fragments (peak II) were produced only during incubation with the *hpa* gene containing viruses. There was no degradation of the HSPG substrate by the culture medium of cells infected with control viruses.

FIG. 4 presents size fractionation of heparanase activity expressed by pF*hpa*2 infected cells. Culture medium of pF*hpa*2 infected High Five cells was applied onto a 50 kDa cut-off membrane. Heparanase activity (conversion of the peak I substrate, (⟡) into peak II HS degradation fragments) was found in the high (> 50 kDa) (•), but not low (< 50 kDa) (o) molecular weight compartment.

FIGs. 5a-b demonstrate the effect of heparin on heparanase activity expressed by pF*hpa*2 and pF*hpa*4 infected High Five cells. Culture media of pF*hpa*2 (5a) and pF*hpa4* (5b) infected High Five cells were incubated (18 h, 37 °C) with sulfate labeled ECM-derived soluble HSPG (peak I, ⟡) in the absence (•) or presence (Δ) of 10 μg/ml heparin. Production of low molecular weight HS degradation fragments was completely abolished in the presence of heparin, a potent inhibitor of heparanase activity (6, 7).

FIGs. 6a-b demonstrate degradation of sulfate labeled intact ECM by virus infected High Five and Sf21 cells. High Five (6a) and Sf21 (6b) cells were plated on sulfate labeled ECM and infected (48 h, 28 °C) with pF*hpa*4 (•) or control pF1 (□) viruses. Control non-infected Sf21 cells (R) were plated on the labeled ECM as well. The pH of the cultured medium was adjusted to 6.0 - 6.2 followed by 24 h incubation at 37 °C. Sulfate labeled material released into the incubation medium was analyzed by gel filtration on Sepharose 6B. HS degradation fragments were produced only by cells infected with the *hpa* containing virus.

FIG. 7a-b demonstrate degradation of sulfate labeled intact ECM by virus infected cells. High Five (7a) and Sf21 (7b) cells were plated on sulfate labeled ECM and infected (48 h, 28 °C) with pF*hpa*4 (•) or control pF1 (□) viruses. Control non-infected Sf21 cells (R) were plate on labeled ECM as well. The pH of the cultured medium was adjusted to 6.0 - 6.2, followed by 48 h incubation at 28 °C. Sulfate labeled degradation fragments released into the incubation

medium was analyzed by gel filtration on Sepharose 6B. HS degradation fragments were produced only by cells infected with the *hpa* containing virus.

FIGs. 8a-b demonstrate degradation of sulfate labeled intact ECM by the culture medium of pF*hpa4* infected cells. Culture media of High Five (8a) and Sf21 (8b) cells that were infected with pF*hpa*4 (•) or control pF1 (□) viruses were incubated (48 h, 37 °C, pH 6.0) with intact sulfate labeled ECM. The ECM was also incubated with the culture medium of control non-infected Sf21 cells (R). Sulfate labeled material released into the reaction mixture was subjected to gel filtration analysis. Heparanase activity was detected only in the culture medium of pF*hpa*4 infected cells.

FIGs. 9a-b demonstrate the effect of heparin on heparanase activity in the culture medium of pF*hpa*4 infected cells. Sulfate labeled ECM was incubated (24 h, 37 °C, pH 6.0) with culture medium of pF*hpa*4 infected High Five (9a) and Sf21 (9b) cells in the absence (•) or presence (V) of 10 μg/ml heparin. Sulfate labeled material released into the incubation medium was subjected to gel filtration on Sepharose 6B. Heparanase activity (production of peak II HS degradation fragments) was completely inhibited in the presence of heparin.

FIGs. 10a-b demonstrate purification of recombinant heparanase on heparin-Sepharose. Culture medium of Sf21 cells infected with pF*hpa*4 virus was subjected to heparin-Sepharose chromatography. Elution of fractions was performed with 0.35 - 2 M NaCl gradient (◇). Heparanase activity in the eluted fractions is demonstrated in Figure 10a (•). Fractions 15-28 were subjected to 15% SDS-polyacrylamide gel electrophoresis followed by silver nitrate staining. A correlation is demonstrated between a major protein band (MW ~ 63,000) in fractions 19 - 24 and heparanase activity.

FIGs. 11a-b demonstrate purification of recombinant heparanase on a Superdex 75 gel filtration column. Active fractions eluted from heparin-Sepharose (Figure 10a) were pooled, concentrated and applied onto Superdex 75 FPLC column. Fractions were collected and aliquots of each fraction were tested for heparanase activity (C, Figure 11a) and analyzed by SDS-polyacrylamide gel electrophoresis followed by silver nitrate staining (Figure 11b). A correlation is seen between the appearance of a major protein band (MW ~ 63,000) in fractions 4 - 7 and heparanase activity.

FIGs. 12a-e demonstrate expression of the *hpa* gene by RT-PCR with total RNA from human embryonal tissues (12a), human extra-embryonal tissues (12b) and cell lines from different origins (12c-e). RT-PCR products using *hpa* specific primers (I), primers for GAPDH housekeeping gene (II), and control reactions without reverse transcriptase demonstrating absence of genomic DNA or other

contamination in RNA samples (III). M- DNA molecular weight marker VI (Boehringer Mannheim). For 12a: lane 1 - neutrophil cells (adult), lane 2 - muscle, lane 3 - thymus, lane 4 - heart, lane 5 - adrenal. For 12b: lane 1 - kidney, lane 2 - placenta (8 weeks), lane 3 - placenta (11 weeks), lanes 4-7 - mole (complete hydatidiform mole), lane 8 - cytotrophoblast cells (freshly isolated), lane 9 - cytotrophoblast cells (1.5 h *in vitro*), lane 10 - cytotrophoblast cells (6 h *in vitro*), lane 11 - cytotrophoblast cells (18 h *in vitro*), lane 12 - cytotrophoblast cells (48 h *in vitro*). For 12c: lane 1 - JAR bladder cell line, lane 2 - NCITT testicular tumor cell line, lane 3 - SW-480 human hepatoma cell line, lane 4 - HTR (cytotrophoblasts transformed by SV40), lane 5 - HPTLP-I hepatocellular carcinoma cell line, lane 6 - EJ-28 bladder carcinoma cell line. For 12d: lane 1 - SK-hep-1 human hepatoma cell line, lane 2 - DAMI human megakaryocytic cell line, lane 3 - DAMI cell line + PMA, lane 4 - CHRF cell line + PMA, lane 5 - CHRF cell line. For 12e: lane 1 - ABAE bovine aortic endothelial cells, lane 2 - 1063 human ovarian cell line, lane 3 - human breast carcinoma MDA435 cell line, lane 4 - human breast carcinoma MDA231 cell line.

FIG. 13 presents a comparison between nucleotide sequences of the human *hpa* and a mouse EST cDNA fragment (SEQ ID NO:12) which is 80 % homologous to the 3' end (starting at nucleotide 1066 of SEQ ID NO:9) of the human *hpa*. The aligned termination codons are underlined.

FIG. 14 demonstrates the chromosomal localization of the *hpa* gene. PCR products of DNA derived from somatic cell hybrids and of genomic DNA of hamster, mouse and human of were separated on 0.7 % agarose gel following amplification with *hpa* specific primers. Lane 1 – Lambda DNA digested with *Bst*EII, lane 2 – no DNA control, lanes 3 – 29, PCR amplification products. Lanes 3-5 – human, mouse and hamster genomic DNA, respectively. Lanes 6-29, human monochromosomal somatic cell hybrids representing chromosomes 1-22 and X and Y, respectively. Lane 30 – Lambda DNA digested with *Bst*EII. An amplification product of approximately 2.8 Kb is observed only in lanes 5 and 9, representing human genomic DNA and DNA derived from cell hybrid carrying human chromosome 4, respectively. These results demonstrate that the *hpa* gene is localized in human chromosome 4.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is of a polynucleotide, referred to hereinbelow interchangeably as *hpa*, *hpa* cDNA or *hpa* gene, encoding a polypeptide having

heparanase activity, vectors including same, transduced cells expressing heparanase and a recombinant protein having heparanase activity.

Before explaining at least one embodiment of the invention in detail, it is to be understood that the invention is not limited in its application to the details of construction and the arrangement of the components set forth in the following description or illustrated in the drawings. The invention is capable of other embodiments or of being practiced or carried out in various ways. Also, it is to be understood that the phraseology and terminology employed herein is for the purpose of description and should not be regarded as limiting.

The present invention can be used to develop treatments for various diseases, to develop diagnostic assays for these diseases and to provide new tools for basic research especially in the fields of medicine and biology.

Specifically, the present invention can be used to develop new drugs to inhibit tumor cell metastasis, inflammation and autoimmunity. The identification of the *hpa* gene encoding for the heparanase enzyme enables the production of a recombinant enzyme in heterologous expression systems.

Furthermore, the present invention can be used to modulate bioavailability of heparin-binding growth factors, cellular responses to heparin-binding growth factors (e.g., bFGF, VEGF) and cytokines (IL-8), cell interaction with plasma lipoproteins, cellular susceptibility to viral, protozoa and some bacterial infections, and disintegration of neurodegenerative plaques. Recombinant heparanase is thus a potential treatment for wound healing, angiogenesis, restenosis, atherosclerosis, inflammation, neurodegenerative diseases (such as, for example, Genstmann-Straussler Syndrome, Creutzfeldt-Jakob disease, Scrape and Alzheimer's disease) and certain viral and some bacterial and protozoa infections. Recombinant heparanase can be used to neutralize plasma heparin, as a potential replacement of protamine.

As used herein, the term "modulate" includes substantially inhibiting, slowing or reversing the progression of a disease, substantially ameliorating clinical symptoms of a disease or condition, or substantially preventing the appearance of clinical symptoms of a disease or condition. A "modulator" therefore includes an agent which may modulate a disease or condition. Modulation of viral, protozoa and bacterial infections includes any effect which substantially interrupts, prevents or reduces any viral, bacterial or protozoa activity and/or stage of the virus, bacterium or protozoon life cycle, or which reduces or prevents infection by the virus, bacterium or protozoon in a subject, such as a human or lower animal.

As used herein, the term "wound" includes any injury to any portion of the body of a subject including, but not limited to, acute conditions such as thermal burns, chemical burns, radiation burns, burns caused by excess exposure to ultraviolet radiation such as sunburn, damage to bodily tissues such as the perineum as a result of labor and childbirth, including injuries sustained during medical procedures such as episiotomies, trauma-induced injuries including cuts, those injuries sustained in automobile and other mechanical accidents, and those caused by bullets, knives and other weapons, and post-surgical injuries, as well as chronic conditions such as pressure sores, bedsores, conditions related to diabetes and poor circulation, and all types of acne, etc.

Anti-heparanase antibodies, raised against the recombinant enzyme, would be useful for immunodetection and diagnosis of micrometastases, autoimmune lesions and renal failure in biopsy specimens, plasma samples, and body fluids. Such antibodies may also serve as neutralizing agents for heparanase activity.

Cloning of the human *hpa* gene encoding heparanase and expressing recombinant heparanase by transfected cells is herein reported. This is the first mammalian heparanase gene to be cloned.

A purified preparation of heparanase isolated from human hepatoma cells was subjected to tryptic digestion and microsequencing.

The YGPDVGQPR (SEQ ID NO:8) sequence revealed was used to screen EST databases for homology to the corresponding back translated DNA sequences. Two closely related EST sequences were identified and were thereafter found to be identical.

Both clones contained an insert of 1020 bp which includes an open reading frame of 973 bp followed by a 3' untranslated region of 27 bp and a Poly A tail, whereas a translation start site was not identified.

Cloning of the missing 5' end was performed by PCR amplification of DNA from placenta Marathon RACE cDNA composite using primers selected according to the EST clones sequence and the linkers of the composite.

A 900 bp PCR fragment, partially overlapping with the identified 3' encoding EST clones was obtained. The joined cDNA fragment (*hpa*), 1721 bp long (SEQ ID NO:9), contained an open reading frame which encodes, as shown in Figure 1 and SEQ ID NO:11, a polypeptide of 543 amino acids (SEQ ID NO:10) with a calculated molecular weight of 61,192 daltons.

A single nucleotide difference at position 799 (A to T) between the EST clones and the PCR amplified cDNA was observed. This difference results in a single amino acid substitution (Tyr to Phe) (Figure 1). Furthermore, the published EST sequences contained an unidentified nucleotide, which following

DNA sequencing of both the EST clones was resolved into two nucleotides (G and C at positions 1630 and 1631 in SEQ ID NO:9, respectively).

The ability of the *hpa* gene product to catalyze degradation of heparan sulfate in an *in vitro* assay was examined by expressing the entire open reading frame in insect cells, using the Baculovirus expression system.

Extracts and conditioned media of cells infected with virus containing the *hpa* gene, demonstrated a high level of heparan sulfate degradation activity both towards soluble ECM-derived HSPG and intact ECM, which was inhibited by heparin, while cells infected with a similar construct containing no *hpa* gene had no such activity, nor did non-infected cells.

The expression pattern of hpa RNA in various tissues and cell lines was investigated using RT-PCR. It was found to be expressed only in tissues and cells previously known to have heparanase activity.

Cloning an extended 5' sequence was enabled from the human SK-hep1 cell line by PCR amplification using the Marathon RACE. The 5' extended sequence of the SK-hep1 *hpa* cDNA was assembled with the sequence of the *hpa* cDNA isolated from human placenta (SEQ ID NO:9). The assembled sequence contained an open reading frame, SEQ ID NOs: 13 and 15, which encodes, as shown in SEQ ID NOs:14 and 15, a polypeptide of 592 amino acids, with a calculated molecular weight of 66,407 daltons. This open reading frame was shown to direct the expression of catalitically active heparanase in a mammalian cell expression system. The expressed heparanase was detectable by anti heparanase antibodies in Western blot analysis.

A panel of monochromosomal human/CHO and human/mouse somatic cell hybrids was used to localize the human heparanase gene to human chromosome 4. The newly isolated heparanase sequence can therefore be used to identify a chromosome region harboring a human heparanase gene in a chromosome spread.

Thus, according to the present invention there is provided a polynucleotide fragment (either DNA or RNA, either single stranded or double stranded) which includes a polynucleotide sequence encoding a polypeptide having heparanase catalytic activity.

The term "heparanase catalytic activity" or its equivalent term "heparanase activity" both refer to a mammalian endoglycosidase hydrolyzing activity which is specific for heparan or heparan sulfate proteoglycan substrates, as opposed to the activity of bacterial enzymes (heparinase I, II and III) which degrade heparin or heparan sulfate by means of β-elimination (37).

In a preferred embodiment of the invention the polynucleotide fragment includes nucleotides 63-1691 of SEQ ID NO:9, or nucleotides 139-1869 of SEQ ID NO:13, which encode the entire human heparanase enzyme.

However, the scope of the present invention is not limited to human heparanase since this is the first disclosure of an open reading frame (ORF) encoding any mammalian heparanase. Using the *hpa* cDNA, parts thereof or synthetic oligonucleotides designed according to its sequence will enable one ordinarily skilled in the art to identify genomic and/or cDNA clones including homologous sequences from other mammalian species.

The present invention is therefore further directed at a polynucleotide fragment which includes a polynucleotide sequence capable of hybridizing (base pairing under either stringent or permissive hybridization conditions, as for example described in Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York.) with *hpa* cDNA, especially with nucleotides 1-721 of SEQ ID NO:9.

In fact, any polynucleotide sequence which encodes a polypeptide having heparanase activity and which shares at least 60 % homology, preferably at least 70 % homology, more preferably at least 80 % homology, most preferably at least 90 % homology with SEQ ID NOs:9 or 13 is within the scope of the present invention.

The polynucleotide fragment according to the present invention may include any part of SEQ ID NOs: 9 or 13. For example, it may include nucleotides 63-721 of SEQ ID NO:9, which is a novel sequence. However, it may include any segment of SEQ ID NOs:9 or 13 which encodes a polypeptide having the heparanase catalytic activity.

When the phrase "encodes a polypeptide having heparanase catalytic activity" is used herein and in the claims section below it refers to the ability of directing the synthesis of a polypeptide which, if so required for its activity, following post translational modifications, such as but not limited to, proteolysis (e.g., removal of a signal peptide and of a pro- or preprotein sequence), methionine modification, glycosylation, alkylation (e.g., methylation), acetylation, etc., is catalytically active in degradation of, for example, ECM and cell surface associated HS.

In a preferred embodiment of the invention the polypeptide encoded by the polynucleotide fragment includes an amino acid sequence as set forth in SEQ ID NOs:10 or 14 or a functional part thereof, i.e., a portion harboring heparanase catalytic activity.

However, any polynucleotide fragment which encodes a polypeptide having heparanase activity is within the scope of the present invention. Therefore, the polypeptide may be allelic, species and/or induced variant of the amino acid sequence set forth in SEQ ID NOs:10 or 14 or functional part thereof.

In fact, any polynucleotide sequence which encodes a polypeptide having heparanase activity, which shares at least 60 % homology, preferably at least 70 % homology, more preferably at least 80 % homology, most preferably at least 90 % homology with SEQ ID NOs:10 or 14 is within the scope of the present invention.

The invention is also directed at providing a single stranded polynucleotide fragment which includes a polynucleotide sequence complementary to at least a portion of a polynucleotide strand encoding a polypeptide having heparanase catalytic activity as described above. The term "complementary" as used herein refers to the ability of base pairing.

The single stranded polynucleotide fragment may be DNA or RNA or even include nucleotide analogs (e.g., thioated nucleotides), it may be a synthetic oligonucleotide or manufactured by transduced host cells, it may be of any desired length which still provides specific base pairing (e.g., 8 or 10, preferably more, nucleotides long) and it may include mismatches that do not hamper base pairing.

The invention is further directed at providing a vector which includes a polynucleotide sequence encoding a polypeptide having heparanase catalytic activity.

The vector may be of any type. It may be a phage which infects bacteria or a virus which infects eukaryotic cells. It may also be a plasmid, phagemid, cosmid, bacmid or an artificial chromosome. The polynucleotide sequence encoding a polypeptide having heparanase catalytic activity may include any of the above described polynucleotide fragments.

The invention is further directed at providing a host cell which includes an exogenous polynucleotide fragment encoding a polypeptide having heparanase catalytic activity.

The exogenous polynucleotide fragment may be any of the above described fragments. The host cell may be of any type. It may be a prokaryotic cell, an eukaryotic cell, a cell line, or a cell as a portion of an organism. The exogenous polynucleotide fragment may be permanently or transiently present in the cell. In other words, transduced cells obtained following stable or transient transfection, transformation or transduction are all within the scope of the present invention. The term "exogenous" as used herein refers to the fact that the

polynucleotide fragment is externally introduced into the cell. Therein it may be present in a single of any number of copies, it may be integrated into one or more chromosomes at any location or be present as an extrachromosomal material.

The invention is further directed at providing a heparanase overexpression system which includes a cell overexpressing heparanase catalytic activity. The cell may be a host cell transiently or stably transfected or transformed with any suitable vector which includes a polynucleotide sequence encoding a polypeptide having heparanase activity and a suitable promoter and enhancer sequences to direct overexpression of heparanase. However, the overexpressing cell may also be a product of an insertion (e.g., via homologous recombination) of a promoter and/or enhancer sequence downstream to the endogenous heparanase gene of the expressing cell, which will direct overexpression from the endogenous gene. The term "overexpression" as used herein in the specification and claims below refers to a level of expression which is higher than a basal level of expression typically characterizing a given cell under otherwise identical conditions.

The invention is further directed at providing a recombinant protein including a polypeptide having heparanase catalytic activity.

The recombinant protein may be purified by any conventional protein purification procedure close to homogeneity and/or be mixed with additives. The recombinant protein may be manufactured using any of the cells described above. The recombinant protein may be in any form. It may be in a crystallized form, a dehydrated powder form or in solution. The recombinant protein may be useful in obtaining pure heparanase, which in turn may be useful in eliciting anti-heparanase antibodies, either poly or monoclonal antibodies, and as a screening active ingredient in an anti-heparanase inhibitors or drugs screening assay or system.

The invention is further directed at providing a pharmaceutical composition which include as an active ingredient a recombinant protein having heparanase catalytic activity.

Formulations for topical administration may include, but are not limited to, lotions, ointments, gels, creams, suppositories, drops, liquids, sprays and powders. Conventional pharmaceutical carriers, aqueous, powder or oily bases, thickeners and the like may be necessary or desirable. Coated condoms, stents, active pads, and other medical devices may also be useful. In fact the scope of the present invention includes any medical equipment such as a medical device containing, as an active ingredient, a recombinant protein having heparanase catalytic activity.

Compositions for oral administration include powders or granules, suspensions or solutions in water or non-aqueous media, sachets, capsules or tablets. Thickeners, diluents, flavorings, dispersing aids, emulsifiers or binders may be desirable.

Formulations for parenteral administration may include, but are not limited to, sterile aqueous solutions which may also contain buffers, diluents and other suitable additives.

Dosing is dependent on severity and responsiveness of the condition to be treated, but will normally be one or more doses per day, with course of treatment lasting from several days to several months or until a cure is effected or a diminution of disease state is achieved. Persons ordinarily skilled in the art can easily determine optimum dosages, dosing methodologies and repetition rates.

Further according to the present invention there is provided a method of identifying a chromosome region harboring a human heparanase gene in a chromosome spread. the method is executed implementing the following method steps, in which in a first step the chromosome spread (either interphase or metaphase spread) is hybridized with a tagged polynucleotyde probe encoding heparanase. The tag is preferably a fluorescent tag. In a second step according to the method the chromosome spread is washed, thereby excess of non-hybridized probe is removed. Finally, signals associated with the hybridized tagged polynucleotyde probe are searched for, wherein detected signals being indicative of a chromosome region harboring the human heparanase gene. One ordinarily skilled in the art would know how to use the sequences disclosed herein in suitable labeling reactions and how to use the tagged probes to detect, using *in situ* hybridization, a chromosome region harboring a human heparanase gene.

Reference is now made to the following examples, which together with the above descriptions, illustrate the invention in a non-limiting fashion.

EXAMPLES

The following protocols and experimental details are referenced in the Examples that follow:

Purification and characterization of heparanase from a human hepatoma cell line and human placenta: A human hepatoma cell line (Sk-hep-1) was chosen as a source for purification of a human tumor-derived heparanase. Purification was essentially as described in U.S. Pat. No. 5,362,641 to Fuks,

which is incorporated by reference as if fully set forth herein. Briefly, 500 liter, $5x10^{11}$ cells were grown in suspension and the heparanase enzyme was purified about 240,000 fold by applying the following steps: (i) cation exchange (CM-Sephadex) chromatography performed at pH 6.0, 0.3-1.4 M NaCl gradient; (ii) cation exchange (CM-Sephadex) chromatography performed at pH 7.4 in the presence of 0.1% CHAPS, 0.3-1.1 M NaCl gradient; (iii) heparin-Sepharose chromatography performed at pH 7.4 in the presence of 0.1% CHAPS, 0.35-1.1 M NaCl gradient; (iv) ConA-Sepharose chromatography performed at pH 6.0 in buffer containing 0.1 % CHAPS and 1 M NaCl, elution with 0.25 M α-methyl mannoside; and (v) HPLC cation exchange (Mono-S) chromatography performed at pH 7.4 in the presence of 0.1 % CHAPS, 0.25-1 M NaCl gradient.

Active fractions were pooled, precipitated with TCA and the precipitate subjected to SDS polyacrylamide gel electrophoresis and/or tryptic digestion and reverse phase HPLC. Tryptic peptides of the purified protein were separated by reverse phase HPLC (C8 column) and homogeneous peaks were subjected to amino acid sequence analysis.

The purified enzyme was applied to reverse phase HPLC and subjected to N-terminal amino acid sequencing using the amino acid sequencer (Applied Biosystems).

Cells: Cultures of bovine corneal endothelial cells (BCECs) were established from steer eyes as previously described (19, 38). Stock cultures were maintained in DMEM (1 g glucose/liter) supplemented with 10 % newborn calf serum and 5 % FCS. bFGF (1 ng/ml) was added every other day during the phase of active cell growth (13, 14).

Preparation of dishes coated with ECM: BCECs (second to fifth passage) were plated into 4-well plates at an initial density of 2 x 10^5 cells/ml, and cultured in sulfate-free Fisher medium plus 5 % dextran T-40 for 12 days. $Na_2{}^{35}SO_4$ (25 μCi/ml) was added on day 1 and 5 after seeding and the cultures were incubated with the label without medium change. The subendothelial ECM was exposed by dissolving (5 min., room temperature) the cell layer with PBS containing 0.5 % Triton X-100 and 20 mM NH_4OH, followed by four washes with PBS. The ECM remained intact, free of cellular debris and firmly attached to the entire area of the tissue culture dish (19, 22).

To prepare soluble sulfate labeled proteoglycans (peak I material), the ECM was digested with trypsin (25 μg/ml, 6 h, 37 °C), the digest was concentrated by reverse dialysis and the concentrated material was applied onto a Sepharose 6B gel filtration column. The resulting high molecular weight

material (Kav< 0.2, peak I) was collected. More than 80 % of the labeled material was shown to be composed of heparan sulfate proteoglycans (11, 39).

Heparanase activity: Cells (1 x 10^6/35-mm dish), cell lysates or conditioned media were incubated on top of ^{35}S-labeled ECM (18 h, 37 °C) in the presence of 20 mM phosphate buffer (pH 6.2). Cell lysates and conditioned media were also incubated with sulfate labeled peak I material (10-20 μl). The incubation medium was collected, centrifuged (18,000 x *g*, 4 °C, 3 min.), and sulfate labeled material analyzed by gel filtration on a Sepharose CL-6B column (0.9 x 30 cm). Fractions (0.2 ml) were eluted with PBS at a flow rate of 5 ml/h and counted for radioactivity using Bio-fluor scintillation fluid. The excluded volume (V_o) was marked by blue dextran and the total included volume (V_t) by phenol red. The latter was shown to comigrate with free sulfate (7, 11, 23). Degradation fragments of HS side chains were eluted from Sepharose 6B at 0.5 < Kav < 0.8 (peak II) (7, 11, 23). A nearly intact HSPG released from ECM by trypsin - and, to a lower extent, during incubation with PBS alone - was eluted next to V_o (Kav < 0.2, peak I). Recoveries of labeled material applied on the columns ranged from 85 to 95 % in different experiments (11). Each experiment was performed at least three times and the variation of elution positions (Kav values) did not exceed +/- 15 %.

Cloning of hpa cDNA: cDNA clones 257548 and 260138 were obtained from the I.M.A.G.E Consortium (2130 Memorial Parkway SW, Hunstville, AL 35801). The cDNAs were originally cloned in *Eco*RI and *Not*I cloning sites in the plasmid vector pT3T7D-Pac. Although these clones are reported to be somewhat different, DNA sequencing demonstrated that these clones are identical to one another. Marathon RACE (rapid amplification of cDNA ends) human placenta (poly-A) cDNA composite was a gift of Prof. Yossi Shiloh of Tel Aviv University. This composite is vector free, as it includes reverse transcribed cDNA fragments to which double, partially single stranded adapters are attached on both sides. The construction of the specific composite employed is described in reference 39a.

Amplification of hp3 PCR fragment was performed according to the protocol provided by Clontech laboratories. The template used for amplification was a sample taken from the above composite. The primers used for amplification were:

First step: 5'-primer: AP1: 5'-CCATCCTAATACGACTCACTATAGGG C-3', SEQ ID NO:1; 3'-primer: HPL229: 5'-GTAGTGATGCCATGTAACTGA ATC-3', SEQ ID NO:2.

Second step: nested 5'-primer: AP2: 5'-ACTCACTATAGGGCTCGAGCG GC-3', SEQ ID NO:3; nested 3'- primer: HPL171: 5'-GCATCTTAGCCGTCT TTCTTCG-3', SEQ ID NO:4. The HPL229 and HPL171 were selected according to the sequence of the EST clones. They include nucleotides 933-956 and 876-897 of SEQ ID NO:9, respectively.

PCR program was 94 °C - 4 min., followed by 30 cycles of 94 °C - 40 sec., 62 °C - 1 min., 72 °C - 2.5 min. Amplification was performed with Expand High Fidelity (Boehringer Mannheim). The resulting ca. 900 bp hp3 PCR product was digested with *Bfr*I and *Pvu*II. Clone 257548 (p*hpa*1) was digested with *Eco*RI, followed by end filling and was then further digested with *Bfr*I. Thereafter the *Pvu*II - *Bfr*I fragment of the hp3 PCR product was cloned into the blunt end - *Bfr*I end of clone p*hpa*1 which resulted in having the entire cDNA cloned in pT3T7-pac vector, designated p*hpa*2.

DNA Sequencing: Sequence determinations were performed with vector specific and gene specific primers, using an automated DNA sequencer (Applied Biosystems, model 373A). Each nucleotide was read from at least two independent primers.

Computer analysis of sequences: Database searches for sequence similarities were performed using the Blast network service. Sequence analysis and alignment of DNA and protein sequences were done using the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the University of Wisconsin.

RT-PCR: RNA was prepared using TRI-Reagent (Molecular research center Inc.) according to the manufacturer instructions. 1.25 μg were taken for reverse transcription reaction using MuMLV Reverse transcriptase (Gibco BRL) and Oligo $(dT)_{15}$ primer, SEQ ID NO:5, (Promega). Amplification of the resultant first strand cDNA was performed with *Taq* polymerase (Promega). The following primers were used:

HPU-355: 5'-TTCGATCCCAAGAAGGAATCAAC-3', SEQ ID NO:6, nucleotides 372-394 in SEQ ID NO:9 or 11.

HPL-229: 5'-GTAGTGATGCCATGTAACTGAATC-3', SEQ ID NO:7, nucleotides 933-956 in SEQ ID NO:9 or 11.

PCR program: 94 °C - 4 min., followed by 30 cycles of 94 °C - 40 sec., 62 °C - 1 min., 72 °C - 1 min.

Expression of recombinant heparanase in insect cells: Cells, High Five and Sf21 insect cell lines were maintained as monolayer cultures in SF900II-SFM medium (GibcoBRL).

Recombinant Baculovirus: Recombinant virus containing the *hpa* gene was constructed using the Bac to Bac system (GibcoBRL). The transfer vector pFastBac was digested with *Sal*I and *Not*I and ligated with a 1.7 kb fragment of p*hpa*2 digested with *Xho*I and *Not*I. The resulting plasmid was designated pFast*hpa*2. An identical plasmid designated pFast*hpa*4 was prepared as a duplicate and both independently served for further experimentations. Recombinant bacmid was generated according to the instructions of the manufacturer with pFast*hpa*2, pFast*hpa*4 and with pFastBac. The latter served as a negative control. Recombinant bacmid DNAs were transfected into Sf21 insect cells. Five days after transfection recombinant viruses were harvested and used to infect High Five insect cells, 3 x 10^6 cells in T-25 flasks. Cells were harvested 2 - 3 days after infection. 4 x 10^6 cells were centrifuged and resuspended in a reaction buffer containing 20 mM phosphate citrate buffer, 50 mM NaCl. Cells underwent three cycles of freeze and thaw and lysates were stored at -80 °C. Conditioned medium was stored at 4 °C.

Partial purification of recombinant heparanase: Partial purification of recombinant heparanase was performed by heparin-Sepharose column chromatography followed by Superdex 75 column gel filtration. Culture medium (150 ml) of Sf21 cells infected with pFhpa4 virus was subjected to heparin-Sepharose chromatography. Elution of 1 ml fractions was performed with 0.35 - 2 M NaCl gradient in presence of 0.1 % CHAPS and 1 mM DTT in 10 mM sodium acetate buffer, pH 5.0. A 25 µl sample of each fraction was tested for heparanase activity. Heparanase activity was eluted at the range of 0.65 - 1.1 M NaCl (fractions 18-26, Figure 10a). 5 µl of each fraction was subjected to 15 % SDS-polyacrylamide gel electrophoresis followed by silver nitrate staining. Active fractions eluted from heparin-Sepharose (Figure 10a) were pooled and concentrated (x 6) on YM3 cut-off membrane. 0.5 ml of the concentrated material was applied onto 30 ml Superdex 75 FPLC column equilibrated with 10 mM sodium acetate buffer, pH 5.0, containing 0.8 M NaCl, 1 mM DTT and 0.1 % CHAPS. Fractions (0.56 ml) were collected at a flow rate of 0.75 ml/min. Aliquots of each fraction were tested for heparanase activity and were subjected to SDS-polyacrylamide gel electrophoresis followed by silver nitrate staining (Figure 11b).

EXAMPLE 1

Cloning of the *hpa* gene

Purified fraction of heparanase isolated from human hepatoma cells (SK-hep-1) was subjected to tryptic digestion and microsequencing. EST (Expressed Sequence Tag) databases were screened for homology to the back translated DNA sequences corresponding to the obtained peptides. Two EST sequences (accession Nos. N41349 and N45367) contained a DNA sequence encoding the peptide YGPDVGQPR (SEQ ID NO:8). These two sequences were derived from clones 257548 and 260138 (I.M.A.G.E Consortium) prepared from 8 to 9 weeks placenta cDNA library (Soares). Both clones which were found to be identical contained an insert of 1020 bp which included an open reading frame (ORF) of 973 bp followed by a 3' untranslated region of 27 bp and a Poly A tail. No translation start site (AUG) was identified at the 5' end of these clones.

Cloning of the missing 5' end was performed by PCR amplification of DNA from a placenta Marathon RACE cDNA composite. A 900 bp fragment (designated hp3), partially overlapping with the identified 3' encoding EST clones was obtained.

The joined cDNA fragment, 1721 bp long (SEQ ID NO:9), contained an open reading frame which encodes, as shown in Figure 1 and SEQ ID NO:11, a polypeptide of 543 amino acids (SEQ ID NO:10) with a calculated molecular weight of 61,192 daltons. The 3' end of the partial cDNA inserts contained in clones 257548 and 260138 started at nucleotide G^{721} of SEQ ID NO:9 and Figure 1.

As further shown in Figure 1, there was a single sequence discrepancy between the EST clones and the PCR amplified sequence, which led to an amino acid substitution from Tyr^{246} in the EST to Phe^{246} in the amplified cDNA. The nucleotide sequence of the PCR amplified cDNA fragment was verified from two independent amplification products. The new gene was designated *hpa*.

As stated above, the 3' end of the partial cDNA inserts contained in EST clones 257548 and 260138 started at nucleotide 721 of *hpa* (SEQ ID NO:9). The ability of the *hpa* cDNA to form stable secondary structures, such as stem and loop structures involving nucleotide stretches in the vicinity of position 721 was investigated using computer modeling. It was found that stable stem and loop structures are likely to be formed involving nucleotides 698-724 (SEQ ID NO:9). In addition, a high GC content, up to 70 %, characterizes the 5' end region of the *hpa* gene, as compared to about only 40 % in the 3' region. These findings may explain the immature termination and therefore lack of 5' ends in the EST clones.

To examine the ability of the *hpa* gene product to catalyze degradation of heparan sulfate in an *in vitro* assay the entire open reading frame was expressed in insect cells, using the Baculovirus expression system. Extracts of cells, infected with virus containing the *hpa* gene, demonstrated a high level of heparan sulfate degradation activity, while cells infected with a similar construct containing no *hpa* gene had no such activity, nor did non-infected cells. These results are further demonstrated in the following Examples.

EXAMPLE 2

Degradation of soluble ECM-derived HSPG

Monolayer cultures of High Five cells were infected (72 h, 28 °C) with recombinant Bacoluvirus containing the pFast*hpa* plasmid or with control virus containing an insert free plasmid. The cells were harvested and lysed in heparanase reaction buffer by three cycles of freezing and thawing. The cell lysates were then incubated (18 h, 37 °C) with sulfate labeled, ECM-derived HSPG (peak I), followed by gel filtration analysis (Sepharose 6B) of the reaction mixture.

As shown in Figure 2, the substrate alone included almost entirely high molecular weight (Mr) material eluted next to V_0 (peak I, fractions 5-20, Kav < 0.35). A similar elution pattern was obtained when the HSPG substrate was incubated with lysates of cells that were infected with control virus. In contrast, incubation of the HSPG substrate with lysates of cells infected with the *hpa* containing virus resulted in a complete conversion of the high Mr substrate into low Mr labeled degradation fragments (peak II, fractions 22-35, 0.5 < Kav < 0.75).

Fragments eluted in peak II were shown to be degradation products of heparan sulfate, as they were (i) 5- to 6-fold smaller than intact heparan sulfate side chains (Kav approx. 0.33) released from ECM by treatment with either alkaline borohydride or papain; and (ii) resistant to further digestion with papain or chondroitinase ABC, and susceptible to deamination by nitrous acid (6, 11).

Similar results (not shown) were obtained with Sf21 cells. Again, heparanase activity was detected in cells infected with the *hpa* containing virus (pF*hpa*), but not with control virus (pF). This result was obtained with two independently generated recombinant viruses. Lysates of control not infected High Five cells failed to degrade the HSPG substrate.

In subsequent experiments, the labeled HSPG substrate was incubated with medium conditioned by infected High Five or Sf21 cells.

As shown in Figures 3a-b, heparanase activity, reflected by the conversion of the high Mr peak I substrate into the low Mr peak II which represents HS degradation fragments, was found in the culture medium of cells infected with the pF*hpa*2 or pF*hpa*4 viruses, but not with the control pF1 or pF2 viruses. No heparanase activity was detected in the culture medium of control non-infected High Five or Sf21 cells.

The medium of cells infected with the pF*hpa*4 virus was passed through a 50 kDa cut off membrane to obtain a crude estimation of the molecular weight of the recombinant heparanase enzyme. As demonstrated in Figure 4, all the enzymatic activity was retained in the upper compartment and there was no activity in the flow through (<50 kDa) material. This result is consistent with the expected molecular weight of the *hpa* gene product.

In order to further characterize the *hpa* product the inhibitory effect of heparin, a potent inhibitor of heparanase mediated HS degradation (40) was examined.

As demonstrated in Figures 5a-b, conversion of the peak I substrate into peak II HS degradation fragments was completely abolished in the presence of heparin.

Altogether, these results indicate that the heparanase enzyme is expressed in an active form by insect cells infected with Baculovirus containing the newly identified human *hpa* gene.

EXAMPLE 3
Degradation of HSPG in intact ECM

Next, the ability of intact infected insect cells to degrade HS in intact, naturally produced ECM was investigated. For this purpose, High Five or Sf21 cells were seeded on metabolically sulfate labeled ECM followed by infection (48 h, 28 °C) with either the pF*hpa*4 or control pF2 viruses. The pH of the medium was then adjusted to pH 6.2-6.4 and the cells further incubated with the labeled ECM for another 48 h at 28 °C or 24 h at 37 °C. Sulfate labeled material released into the incubation medium was analyzed by gel filtration on Sepharose 6B.

As shown in Figures 6a-b and 7a-b, incubation of the ECM with cells infected with the control pF2 virus resulted in a constant release of labeled material that consisted almost entirely (>90%) of high Mr fragments (peak I) eluted with or next to V_0. It was previously shown that a proteolytic activity residing in the ECM itself and/or expressed by cells is responsible for release of the high Mr material (6). This nearly intact HSPG provides a soluble substrate

for subsequent degradation by heparanase, as also indicated by the relatively large amount of peak I material accumulating when the heparanase enzyme is inhibited by heparin (6, 7, 12, Figure 9). On the other hand, incubation of the labeled ECM with cells infected with the pF*hpa*4 virus resulted in release of 60-70% of the ECM-associated radioactivity in the form of low Mr sulfate-labeled fragments (peak II, 0.5 <Kav< 0.75), regardless of whether the infected cells were incubated with the ECM at 28 °C or 37 °C. Control intact non-infected Sf21 or High Five cells failed to degrade the ECM HS side chains.

In subsequent experiments, as demonstrated in Figures 8a-b, High Five and Sf21 cells were infected (96 h, 28 °C) with pF*hpa*4 or control pF1 viruses and the culture medium incubated with sulfate-labeled ECM. Low Mr HS degradation fragments were released from the ECM only upon incubation with medium conditioned by pF*hpa*4 infected cells. As shown in Figure 9, production of these fragments was abolished in the presence of heparin. No heparanase activity was detected in the culture medium of control, non-infected cells. These results indicate that the heparanase enzyme expressed by cells infected with the pF*hpa*4 virus is capable of degrading HS when complexed to other macromolecular constituents (i.e. fibronectin, laminin, collagen) of a naturally produced intact ECM, in a manner similar to that reported for highly metastatic tumor cells or activated cells of the immune system (6, 7).

EXAMPLE 4

Purification of recombinant heparanase

The recombinant heparanase was partially purified from medium of pF*hpa*4 infected Sf21 cells by Heparin-Sepharose chromatography (Figure 10a) followed by gel filtration of the pooled active fractions over an FPLC Superdex 75 column (Figure 11a). A ~ 63 kDa protein was observed, whose quantity, as was detected by silver stained SDS-polyacrylamide gel electrophoresis, correlated with heparanase activity in the relevant column fractions (Figures 10b and 11b, respectively). This protein was not detected in the culture medium of cells infected with the control pF1 virus and was subjected to a similar fractionation on heparin-Sepharose (not shown).

EXAMPLE 5

Expression of the *hpa* gene in various cell types, organs and tissues

Referring now to Figures 12a-e, RT-PCR was applied to evaluate the expression of the *hpa* gene by various cell types and tissues. For this purpose, total RNA was reverse transcribed and amplified. The expected 585 bp long

cDNA was clearly demonstrated in human kidney, placenta (8 and 11 weeks) and mole tissues, as well as in freshly isolated and short termed (1.5-48 h) cultured human placental cytotrophoblastic cells (Figure 12a), all known to express a high heparanase activity (41). The *hpa* transcript was also expressed by normal human neutrophils (Figure 12b). In contrast, there was no detectable expression of the *hpa* mRNA in embryonic human muscle tissue, thymus, heart and adrenal (Figure 12b). The *hpa* gene was expressed by several, but not all, human bladder carcinoma cell lines (Figure 12c), SK hepatoma (SK-hep-1), ovarian carcinoma (OV 1063), breast carcinoma (435, 231), melanoma and megakaryocytic (DAMI, CHRF) human cell lines (Figures 12d-e).

The above described expression pattern of the *hpa* transcript was determined to be in a very good correlation with heparanase activity levels determined in various tissues and cell types (not shown).

EXAMPLE 6

hpa homologous genes

EST databases were screened for sequences homologous to the *hpa* gene. Three mouse ESTs were identified (accession No. Aa177901, from mouse spleen, Aa067997 from mouse skin, Aa47943 from mouse embryo), assembled into a 824 bp cDNA fragment which contains a partial open reading frame (lacking a 5' end) of 629 bp and a 3' untranslated region of 195 bp (SEQ ID NO:12). As shown in Figure 13, the coding region is 80% similar to the 3' end of the *hpa* cDNA sequence. These ESTs are probably cDNA fragments of the mouse *hpa* homolog that encodes for the mouse heparanase.

Searching for consensus protein domains revealed an amino terminal homology between the heparanase and several precursor proteins such as Procollagen Alpha 1 precursor, Tyrosine-protein kinase-RYK, Fibulin-1, Insulin-like growth factor binding protein and several others. The amino terminus is highly hydrophobic and contains a potential trans-membrane domain. The homology to known signal peptide sequences suggests that it could function as a signal peptide for protein localization.

EXAMPLE 7

Isolation of an extended 5' end of hpa cDNA from human SK-hep1 cell line

The 5' end of *hpa* cDNA was isolated from human SK-hep1 cell line by PCR amplification using the Marathon RACE (rapid amplification of cDNA

ends) kit (Clontech). Total RNA was prepared from SK-hep1 cells using the TRI-Reagent (Molecular research center Inc.) according to the manufacturer instructions. Poly A+ RNA was isolated using the mRNA separator kit (Clonetech).

The Marahton RACE SK-hep1 cDNA composite was constructed according to the manufacturer recommendations. First round of amplification was performed using an adaptor specific primer AP1: 5'-CCATCCTAATACG ACTCACTATAGGGC-3', SEQ ID NO:1, and a *hpa* specific antisense primer hpl-629: 5'-CCCCAGGAGCAGCAGCATCAG-3', SEQ ID NO:17, corresponding to nucleotides 119-99 of SEQ ID NO:9. The resulting PCR product was subjected to a second round of amplification using an adaptor specific nested primer AP2: 5'-ACTCACTATAGGGCTCGAGCGGC-3', SEQ ID NO:3, and a *hpa* specific antisense nested primer hpl-666 5'-AGGCTTCGAGCGCAGCAGCAT-3', SEQ ID NO:18, corresponding to nucleotides 83-63 of SEQ ID NO:9. The PCR program was as follows: a hot start of 94 °C for 1 minute, followed by 30 cycles of 90 °C - 30 seconds, 68 °C - 4 minutes. The resulting 300 bp DNA fragment was extracted from an agarose gel and cloned into the vector pGEM-T Easy (Promega). The resulting recombinant plasmid was designated pHPSK1.

The nucleotide sequence of the pHPSK1 insert was determined and it was found to contain 62 nucleotides of the 5' end of the placenta *hpa* cDNA (SEQ ID NO:9) and additional 178 nucleotides upstream, the first 178 nucleotides of SEQ ID NOs:13 and 15.

A single nucleotide discrepancy was identified between the SK-hep1 cDNA and the placenta cDNA. The "T" derivative at position 9 of the placenta cDNA (SEQ ID NO:9), is replaced by a "C" derivative at the corresponding position 187 of the SK-hep1 cDNA (SEQ ID NO:13).

The discrepancy is likely to be due to a mutation at the 5' end of the placenta cDNA clone as confirmed by sequence analysis of sevsral additional cDNA clones isolated from placenta, which like the SK-hep1 cDNA contained C at position 9 of SEQ ID NO:9.

The 5' extended sequence of the SK-hep1 *hpa* cDNA was assembled with the sequence of the *hpa* cDNA isolated from human placenta (SEQ ID NO:9). The assembled sequence contained an open reading frame which encodes, as shown in SEQ ID NOs:14 and 15, a polypeptide of 592 amino acids with a calculated molecular weight of 66,407 daltons. The open reading frame is flanked by 93 bp 5' untranslated region (UTR).

EXAMPLE 8

Isolation of the upstream genomic region of the hpa gene

The upstream region of the *hpa* gene was isolated using the Genome Walker kit (Clontech) according to the manufacturer recommendations. The kit includes five human genomic DNA samples each digested with a different restriction endonuclease creating blunt ends: *Eco*RV, *Sca*I, *Dra*I, *Pvu*II and *Ssp*I.

The blunt ended DNA fragments are ligated to partially single stranded adaptors. The Genomic DNA samples were subjected to PCR amplification using the adaptor specific primer and a gene specific primer. Amplification was performed with Expand High Fidelity (Boehringer Mannheim).

A first round of amplification was performed using the ap1 primer: 5'-G TAATACGACTCACTATAGGGC-3', SEQ ID NO:19, and the *hpa* specific antisense primer hpl-666: 5'-AGGCTTCGAGCGCAGCAGCAT-3', SEQ ID NO:18, corresponding to nucleotides 83 – 63 of SEQ ID NO:9. The PCR program was as follows: a hot start of 94 °C - 3 minutes, followed by 36 cycles of 94 °C - 40 seconds, 67 °C - 4 minutes.

The PCR products of the first amplification were diluted 1:50. One μl of the diluted sample was used as a template for a second amplification using a nested adaptor specific primer ap2: 5'-ACTATAGGGCACGCGTGGT-3', SEQ ID NO:20, and a *hpa* specific antisense primer hpl-690, 5'-CTTGGGCTCACC TGGCTGCTC-3', SEQ ID NO:21, corresponding to nucleotides 62-42 of SEQ ID NO:9. The resulting amplification products were analyzed using agarose gel electrophoresis. Five different PCR products were obtained from the five amplification reactions. A DNA fragment of approximately 750 bp which was obtained from the *Ssp*I digested DNA sample was gel extracted. The purified fragment was ligated into the plasmid vector pGEM-T Easy (Promega). The resulting recombinant plasmid was designated pGHP6905 and the nucleotide sequence of the *hpa* insert was determined.

A partial sequence of 594 nucleotides is shown in SEQ ID NO:16. The last nucleotide in SEQ ID NO:13 corresponds to nucleotide 93 in SEQ ID:13. The DNA sequence in SEQ ID NO:16 contains the 5' region of the *hpa* cDNA and 501 nucleotides of the genomic upstream region which are predicted to contain the promoter region of the *hpa* gene.

EXAMPLE 9

Expression of the 592 amino acids HPA polypeptide in a human 293 cell line

The 592 amino acids open reading frame (SEQ ID NOs:13 and 15) was constructed by ligation of the 110 bp corresponding to the 5' end of the SK-hep1 *hpa* cDNA with the placenta cDNA. More specifically the Marathon RACE - PCR amplification product of the placenta *hpa* DNA was digested with *Sac*I and an approximately 1 kb fragment was ligated into a *Sac*I-digested pGHP6905 plasmid. The resulting plasmid was digested with *Ear*I and *Aat*II. The *Ear*I sticky ends were blunted and an approximately 280 bp *Ear*I/blunt-*Aat*II fragment was isolated. This fragment was ligated with pFast*hpa* digested with *Eco*RI which was blunt ended using Klenow fragment and further digested with *Aat*II. The resulting plasmid contained a 1827 bp insert which includes an open reading frame of 1776 bp, 31 bp of 3' UTR and 21 bp of 5' UTR. This plasmid was designated pFastL*hpa*.

A mammalian expression vector was constructed to drive the expression of the 592 amino acids heparanase polypeptide in human cells. The *hpa* cDNA was excised prom pFastL*hpa* with *Bss*HII and *Not*I. The resulting 1850 bp *Bss*HII-*Not*I fragment was ligated to a mammalian expression vector pSI (Promega) digested with *Mlu*I and *Not*I. The resulting recombinant plasmid, pSI*hpa*Met2 was transfected into a human 293 embryonic kidney cell line.

Transient expression of the 592 amino-acids heparanase was examined by western blot analysis and the enzymatic activity was tested using the gel shift assay. Both these procedures are described in length in U.S. Pat. application No. 09/071,739, filed May 1, 1998, which is incorporated by reference as if fully set forth herein. Cells were harvested 3 days following transfection. Harvested cells were re-suspended in lysis buffer containing 150 mM NaCl, 50 mM Tris pH 7.5, 1% Triton X-100, 1 mM PMSF and protease inhibitor cocktail (Boehringer Mannheim). 40 μg protein extract samples were used for separation on a SDS-PAGE. Proteins were transferred onto a PVDF Hybond-P membrane (Amersham). The membrane was incubated with an affinity purified polyclonal anti heparanase antibody, as described in U.S. Pat. application No. 09/071,739. A major band of approximately 50 kDa was observed in the transfected cells as well as a minor band of approximately 65 kDa. A similar pattern was observed in extracts of cells transfected with the pS*hpa* as demonstrated in U.S. Pat. application No. 09/071,739. These two bands probably represent two forms of the recombinant heparanase protein produced by the transfected cells. The 65 kDa protein probably represents a heparanase precursor, while the 50 kDa protein is suggested herein to be the processed or mature form.

The catalytic activity of the recombinant protein expressed in the pS*hpa*Met2 transfected cells was tested by gel shift assay. Cell extracts of transfected and of mock transfected cells were incubated overnight with heparin (6 μg in each reaction) at 37 °C, in the presence of 20 mM phosphate citrate buffer pH 5.4, 1 mM $CaCl_2$, 1 mM DTT and 50 mM NaCl. Reaction mixtures were then separated on a 10 % polyacrylamide gel. The catalytic activity of the recombinant heparanase was clearly demonstrated by a faster migration of the heparin molecules incubated with the transfected cell extract as compared to the control. Faster migration indicates the disappearance of high molecular weight heparin molecules and the generation of low molecular weight degradation products.

EXAMPLE 10

Chromosomal localization of the hpa gene

Chromosomal mapping of the *hpa* gene was performed utilizing a panel of monochromosomal human/CHO and human/mouse somatic cell hybrids, obtained from the UK HGMP Resource Center (Cambridge, England).

40 ng of each of the somatic cell hybrid DNA samples were subjected to PCR amplification using the *hpa* primers: hpu565 5'-AGCTCTGTAGATGTGC TATACAC-3', SEQ ID NO:22, corresponding to nucleotides 564-586 of SEQ ID NO:9 and an antisense primer hpl171 5'-GCATCTTAGCCGTCTTTCTTCG-3', SEQ ID NO:23, corresponding to nucleotides 897-876 of SEQ ID NO:9.

The PCR program was as follows: a hot start of 94 °C – 3 minutes, followed by 7 cycles of 94 °C – 45 seconds, 66 °C – 1 minute, 68 °C – 5 minutes, followed by 30 cycles of 94 °C – 45 seconds, 62 °C – 1 minute, 68 °C – 5 minutes, and a 10 minutes final extension at 72 °C.

The reactions were performed with Expand long PCR (Boehringer Mannheim). The resulting amplification products were analyzed using agarose gel electrophoresis. As demonstrated in Figure 14, a single band of approximately 2.8 Kb was obtained from chromosome 4, as well as from the control human genomic DNA. A 2.8 kb amplification product is expected based on amplification of the genomic *hpa* clone (data not shown). No amplification products were obtained neither in the control DNA samples of hamster and mouse nor in somatic hybrids of other human chromosome.

Although the invention has been described in conjunction with specific embodiments thereof, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art. Accordingly, it is intended

to embrace all such alternatives, modifications and variations that fall within the spirit and broad scope of the appended claims.

LIST OF REFERENCES CITED HEREINABOVE BY NUMBERS

1. Wight, T.N., Kinsella, M.G., and Qwarnstromn, E.E. (1992). The role of proteoglycans in cell adhesion, migration and proliferation. *Curr. Opin. Cell Biol.,* **4,** 793-801.

2. Jackson, R.L., Busch, S.J., and Cardin, A.L. (1991). Glycosaminoglycans: Molecular properties, protein interactions and role in physiological processes. *Physiol. Rev.,* **71,** 481-539.

3. Wight, T.N. (1989). Cell biology of arterial proteoglycans. *Arteriosclerosis,* **9,** 1-20.

4. Kjellen, L., and Lindahl, U. (1991). Proteoglycans: structures and interactions. *Annu. Rev. Biochem.,* **60,** 443-475.

5. Ruoslahti, E., and Yamaguchi, Y. (1991). Proteoglycans as modulators of growth factor activities. *Cell,* **64,** 867-869.

6. Vlodavsky, I., Eldor, A., Haimovitz-Friedman, A., Matzner, Y., Ishai-Michaeli, R., Levi, E., Bashkin, P., Lider, O., Naparstek, Y., Cohen, I.R., and Fuks, Z. (1992). Expression of heparanase by platelets and circulating cells of the immune system: Possible involvement in diapedesis and extravasation. *Invasion & Metastasis,* **12,** 112-127.

7. Vlodavsky, I., Mohsen, M., Lider, O., Ishai-Michaeli, R., Ekre, H.-P., Svahn, C.M., Vigoda, M., and Peretz, T. (1995). Inhibition of tumor metastasis by heparanase inhibiting species of heparin. *Invasion & Metastasis,* **14,** 290-302.

8. Nakajima, M., Irimura, T., and Nicolson, G.L. (1988). Heparanase and tumor metastasis. *J. Cell. Biochem.,* **36,** 157-167.

9. Nicolson, G.L. (1988). Organ specificity of tumor metastasis: Role of preferential adhesion, invasion and growth of malignant cells at specific secondary sites. *Cancer Met. Rev.,* **7,** 143-188.

10. Liotta, L.A., Rao, C.N., and Barsky, S.H. (1983). Tumor invasion and the extracellular matrix. *Lab. Invest.,* **49,** 639-649.

11. Vlodavsky, I., Fuks, Z., Bar-Ner, M., Ariav, Y., and Schirrmacher, V. (1983). Lymphoma cell mediated degradation of sulfated proteoglycans in the subendothelial extracellular matrix: Relationship to tumor cell metastasis. *Cancer Res.,* **43,** 2704-2711.

12. Vlodavsky, I., Ishai-Michaeli, R., Bar-Ner, M., Fridman, R., Horowitz, A.T., Fuks,Z. and Biran, S. (1988). Involvement of heparanase in tumor metastasis and angiogenesis. Is. *J. Med.,* **24**, 464-470.

13. Vlodavsky, I., Liu, G.M., and Gospodarowicz, D. (1980). Morphological appearance, growth behavior and migratory activity of human tumor cells maintained on extracellular matrix vs. plastic. *Cell,* **19,** 607-616.

14. Gospodarowicz, D., Delgado, D., and Vlodavsky, I. (1980). Permissive effect of the extracellular matrix on cell proliferation in-vitro. *Proc. Natl. Acad. Sci. USA.,* **77,** 4094-4098.

15. Bashkin, P., Doctrow, S., Klagsbrun, M., Svahn, C.M., Folkman, J., and Vlodavsky, I. (1989). Basic fibroblast growth factor binds to subendothelial extracellular matrix and is released by heparitinase and heparin-like molecules. *Biochemistry,* **28,** 1737-1743.

16. Parish, C.R., Coombe, D.R., Jakobsen, K.B., and Underwood, P.A. (1987). Evidence that sulphated polysaccharides inhibit tumor metastasis by blocking tumor cell-derived heparanase. *Int. J. Cancer,* **40,** 511-517.

16a. Vlodavsky, I., Hua-Quan Miao., Benezra, M., Lider, O., Bar-Shavit, R., Schmidt, A., and Peretz, T. (1997). Involvement of the extracellular matrix, heparan sulfate proteoglycans and heparan sulfate degrading enzymes in angiogenesis and metastasis. In: Tumor Angiogenesis. Eds. C.E. Lewis, R. Bicknell & N. Ferrara. Oxford University Press, Oxford UK, pp. 125-140.

17. Burgess, W.H., and Maciag, T. (1989). The heparin-binding (fibroblast) growth factor family of proteins. *Annu. Rev. Biochem.,* **58,** 575-606.

18. Folkman, J., and Klagsbrun, M. (1987). Angiogenic factors. *Science,* **235,** 442-447.

19. Vlodavsky, I., Folkman, J., Sullivan, R., Fridman, R., Ishai-Michaelli, R., Sasse, J., and Klagsbrun, M. (1987). Endothelial cell-derived basic fibroblast growth factor: Synthesis and deposition into subendothelial extracellular matrix. .*Proc. Natl. Acad. Sci. USA,* **84,** 2292-2296.

20. Folkman, J., Klagsbrun, M., Sasse, J., Wadzinski, M., Ingber, D., and Vlodavsky, I. (1980). A heparin-binding angiogenic protein - basic fibroblast growth factor - is stored within basement membrane. *Am. J. Pathol.,* **130,** 393-400.

21. Cardon-Cardo, C., Vlodavsky, I., Haimovitz-Friedman, A., Hicklin, D., and Fuks, Z. (1990). Expression of basic fibroblast growth factor in normal human tissues. *Lab. Invest.,* **63,** 832-840.

22. Ishai-Michaeli, R., Svahn, C.-M., Chajek-Shaul, T., Korner, G., Ekre, H.-P., and Vlodavsky, I. (1992). Importance of size and sulfation of heparin in release of basic fibroblast factor from the vascular endothelium and extracellular matrix. *Biochemistry,* **31,** 2080-2088.

23. Ishai-Michaeli, R., Eldor, A., and Vlodavsky, I. (1990). Heparanase activity expressed by platelets, neutrophils and lymphoma cells releases active fibroblast growth factor from extracellular matrix. *Cell Reg.,* **1,** 833-842.

24. Vlodavsky, I., Bar-Shavit, R., Ishai-Michaeli, R., Bashkin, P., and Fuks, Z. (1991). Extracellular sequestration and release of fibroblast growth factor: a regulatory mechanism? *Trends Biochem. Sci.,* **16,** 268-271.

25. Vlodavsky, I., Bar-Shavit, R., Korner, G., and Fuks, Z. (1993). Extracellular matrix-bound growth factors, enzymes and plasma proteins. In Basement membranes: Cellular and molecular aspects (eds. D.H. Rohrbach and R. Timpl), pp327-343. Academic press Inc., Orlando, Fl.

26. Yayon, A., Klagsbrun, M., Esko, J.D., Leder, P., and Ornitz, D.M. (1991). Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. *Cell,* **64,** 841-848.

27. Spivak-Kroizman, T., Lemmon, M.A., Dikic, I., Ladbury, J.E., Pinchasi, D., Huang, J., Jaye, M., Crumley, G., Schlessinger, J., and Lax, I. (1994). Heparin-induced oligomerization of FGF molecules is responsible for FGF receptor dimerization, activation, and cell proliferation. *Cell,* **79,** 1015-1024.

28. Ornitz, D.M., Herr, A.B., Nilsson, M., West, a., J., Svahn, C.-M., and Waksman, G. (1995). FGF binding and FGF receptor activation by synthetic heparan-derived di- and trisaccharides. *Science,* **268,** 432-436.

29. Gitay-Goren, H., Soker, S., Vlodavsky, I., and Neufeld, G. (1992). Cell surface associated heparin-like molecules are required for the binding of vascular endothelial growth factor (VEGF) to its cell surface receptors. *J. Biol. Chem.,* **267,** 6093-6098.

30. Lider, O., Baharav, E., Mekori, Y., Miller, T., Naparstek, Y., Vlodavsky, I., and Cohen, I.R. (1989). Suppression of experimental autoimmune diseases and prolongation of allograft survival by treatment of animals with heparinoid inhibitors of T lymphocyte heparanase. *J. Clin. Invest.,* **83,** 752-756.

31. Lider, O., Cahalon, L., Gilat, D., Hershkovitz, R., Siegel, D., Margalit, R., Shoseyov, O., and Cohn, I.R. (1995). A disaccharide that inhibits tumor necrosis factor α is formed from the extracellular matrix by the enzyme heparanase. *Proc. Natl. Acad. Sci. USA.,* **92,** 5037-5041.

31a. Rapraeger, A., Krufka, A., and Olwin, B.R. (1991). Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. *Science*, **252,** 1705-1708.

32. Eisenberg, S., Sehayek, E., Olivecrona, T., and Vlodavsky, I. (1992). Lipoprotein lipase enhances binding of lipoproteins to heparan sulfate on cell surfaces and extracellular matrix. *J. Clin. Invest.,* **90,** 2013-2021.

33. Shieh, M-T., Wundunn, D., Montgomery, R.I., Esko, J.D., and Spear, P.G. J. (1992). Cell surface receptors for herpes simplex virus are heparan sulfate proteoglycans. *J Cell Biol.*, **116**, 1273-1281.

33a. Chen, Y., Maguire, T., Hileman, R.E., Fromm, J.R., Esko, J.D., Linhardt, R.J., and Marks, R.M. (1997). Dengue virus infectivity depends on envelope protein binding to target cell heparan sulfate. *Nature Medicine* **3,** 866-871.

33b. Putnak, J.R., Kanesa-Thasan, N., and Innis, B.L. (1997). A putative cellular receptor for dengue viruses. *Nature Medicine* **3,** 828-829.

34. Narindrasorasak, S., Lowery, D., Gonzalez-DeWhitt, P., Poorman, R.A., Greenberg, B., Kisilevsky, R. (1991). High affinity interactions between the Alzheimer's beta-amyloid precursor protein and the basement membrane form of theparan sulfate proteoglycan. *J. Biol. Chem.,* **266,** *12878-83.*

35. Ross, R. (1993). The pathogenesis of atherosclerosis: a perspective for the 1990s. *Nature (Lond.)*., **362**:801-809.

36. Zhong-Sheng, J., Walter, J., Brecht, R., Miranda, D., Mahmood Hussain, M., Innerarity, T.L. and Mahley, W.R. (1993). Role of heparan sulfate proteoglycans in the binding and uptake of apolipoprotein E-enriched remnant lipoproteins by cultured cells. *J. Biol. Chem.*, **268,** 10160-10167.

37. Ernst, S., Langer, R., Cooney, Ch.L., and Sasisekharan, R. (1995). Enzymatic degradation of glycosaminoglycans. Critical Reviews in Biochemistry and Molecular Biology, 30(5), 387-444.

38. Gospodarowicz, D., Mescher, AL., Birdwell, CR. (1977). Stimulation of corneal endothelial cell proliferation in vitro by fibroblast and epidermal growth factors. *Exp Eye Res* **25**, 75-89.

39. Haimovitz-Friedman, A., Falcone, D.J., Eldor, A., Schirrmacher, V.,
Vlodavsky, I., and Fuks, Z. (1991) Activation of platelet heparitinase by tumor cell-derived factors. *Blood*, **78**, 789-796.

39a. Savitsky, K., Platzer, M., Uziel, T., Gilad, S., Sartiel, A., Rosental, A., Elroy-Stein, O., Siloh, Y. and Rotman, G. (1997). Ataxia-telangiectasia: structural diversity of untranslated sequences suggests complex post-translational regulation of ATM gene expression. Nucleic Acids Res. 25(9), 1678-1684.

40. Bar-Ner, M., Eldor, A., Wasserman, L., Matzner, Y., and Vlodavsky, I. (1987). Inhibition of heparanase mediated degradation of extracellular matrix heparan sulfate by modified and non-anticoagulant heparin species. *Blood,* **70,** 551-557.

41. Goshen, R., Hochberg, A., Korner, G., Levi, E., Ishai- Michaeli, R., Elkin, M., de Grot, N., and Vlodavsky, I. (1996). Purification and characterization of placental heparanase and its expression by cultured cytotrophoblasts. *Mol. Human Reprod.* **2,** 679-684.

WHAT IS CLAIMED IS:

1. A polynucleotide fragment comprising a polynucleotide sequence encoding a polypeptide having heparanase catalytic activity.

2. The polynucleotide fragment of claim 1, wherein said polynucleotide sequence includes nucleotides 63-1691 of SEQ ID NO:9, or nucleotides 139-1869 of SEQ ID NO:13.

3. The polynucleotide fragment of claim 1, wherein said polynucleotide sequence includes nucleotides 63-721 of SEQ ID NO:9.

4. The polynucleotide fragment of claim 1, wherein said polynucleotide is as set forth in SEQ ID NOs:9 or 13.

5. The polynucleotide fragment of claim 1, wherein said polynucleotide sequence includes a segment of SEQ ID NOs:9 or 13, said segment encodes said polypeptide having said heparanase catalytic activity.

6. The polynucleotide fragment of claim 1, wherein said polypeptide includes an amino acid sequence as set forth in SEQ ID NOs:10 or 14.

7. The polynucleotide fragment of claim 1, wherein said polypeptide includes a segment of SEQ ID NOs:10 or 14, said segment harbors said heparanase catalytic activity.

8. The polynucleotide fragment of claim 1, wherein said polynucleotide sequence is selected from the group consisting of double stranded DNA, single stranded DNA and RNA.

9. A single stranded polynucleotide fragment comprising a polynucleotide sequence complementary to at least a portion of a polynucleotide strand encoding a polypeptide having heparanase catalytic activity.

10. The polynucleotide fragment of claim 9, wherein said polynucleotide sequence includes at least a portion of SEQ ID NOs:9 or 13.

11. A vector comprising a polynucleotide sequence encoding a polypeptide having heparanase catalytic activity.

12. The vector of claim 11, wherein said polynucleotide sequence includes nucleotides 63-1691 of SEQ ID NO:9, or nucleotides 139-1869 of SEQ ID NO:13.

13. The vector of claim 11, wherein said polynucleotide sequence includes nucleotides 63-721 of SEQ ID NO:9.

14. The vector of claim 11, wherein said polynucleotide sequence is as set forth in SEQ ID NOs:9 or 13.

15. The vector of claim 11, wherein said polynucleotide sequence includes a segment of SEQ ID NOs:9 or 13, said segment encodes said polypeptide having said heparanase catalytic activity.

16. The vector of claim 11, wherein said polypeptide includes an amino acid sequence as set forth in SEQ ID NOs:10 or 14.

17. The vector of claim 11, wherein said polypeptide includes a segment of SEQ ID NOs:10 or 14, said segment harbors said heparanase catalytic activity.

18. The vector of claim 11, wherein said polynucleotide sequence is selected from the group consisting of double stranded DNA, single stranded DNA and RNA.

19. A host cell comprising an exogenous polynucleotide fragment including a polynucleotide sequence encoding a polypeptide having heparanase catalytic activity.

20. The host cell of claim 19, wherein said polynucleotide sequence includes nucleotides 63-1691 of SEQ ID NO:9, or nucleotides 139-1869 of SEQ ID NO:13.

21. The host cell of claim 19, wherein said polynucleotide sequence includes nucleotides 63-721 of SEQ ID NO:9.

22. The host cell of claim 19, wherein said polynucleotide sequence is as set forth in SEQ ID NOs:9 or 13.

23. The host cell of claim 19, wherein said polynucleotide sequence includes a segment of SEQ ID NOs:9 or 13, said segment encodes said polypeptide having said heparanase catalytic activity.

24. The host cell of claim 19, wherein said polypeptide includes an amino acid sequence as set forth in SEQ ID NOs:10 or 14.

25. The host cell of claim 19, wherein said polypeptide includes a segment of SEQ ID NOs:10 or 14, said segment harbors said heparanase catalytic activity.

26. The host cell of claim 19, wherein said polynucleotide sequence is selected from the group consisting of double stranded DNA, single stranded DNA and RNA.

27. A host cell expressing a recombinant heparanase.

28. A recombinant protein comprising a polypeptide having heparanase catalytic activity.

29. The recombinant protein of claim 28, wherein said polypeptide includes a segment of SEQ ID NOs:10 or 14.

30. A polynucleotide fragment comprising a polynucleotide sequence capable of hybridizing with nucleotides 1-721 of SEQ ID NO:9.

31. A polynucleotide sequence as set forth in SEQ ID NOs:9 or 13.

32. A polynucleotide sequence homologous to SEQ ID NOs:9 or 13.

33. An amino acid sequence as set forth in SEQ ID NOs:10 or 14.

34. An amino acid sequence homologous to SEQ ID NOs:10 or 14.

35. A pharmaceutical composition comprising as an active ingredient a recombinant protein having heparanase catalytic activity.

36. A heparanase overexpression system comprising a cell overexpressing heparanase catalytic activity.

37. A modulator of heparin-binding growth factors, cellular responses to heparin-binding growth factors and cytokines, cell interaction with plasma lipoproteins, cellular susceptibility to viral, protozoa and bacterial infections or disintegration of neurodegenerative plaques comprising as an active ingredient a recombinant protein having heparanase catalytic activity.

38. A medical equipment comprising a medical device containing, as an active ingredient, a recombinant protein having heparanase catalytic activity.

39. The vector of claim 11, wherein said vector is a baculovirus vector.

40. The host cell of claim 19, wherein said cell is an insect cell.

41. The host cell of claim 27, wherein said cell is an insect cell.

42. A method of identifying a chromosome region harboring a human heparanase gene in a chromosome spread comprising the steps of:

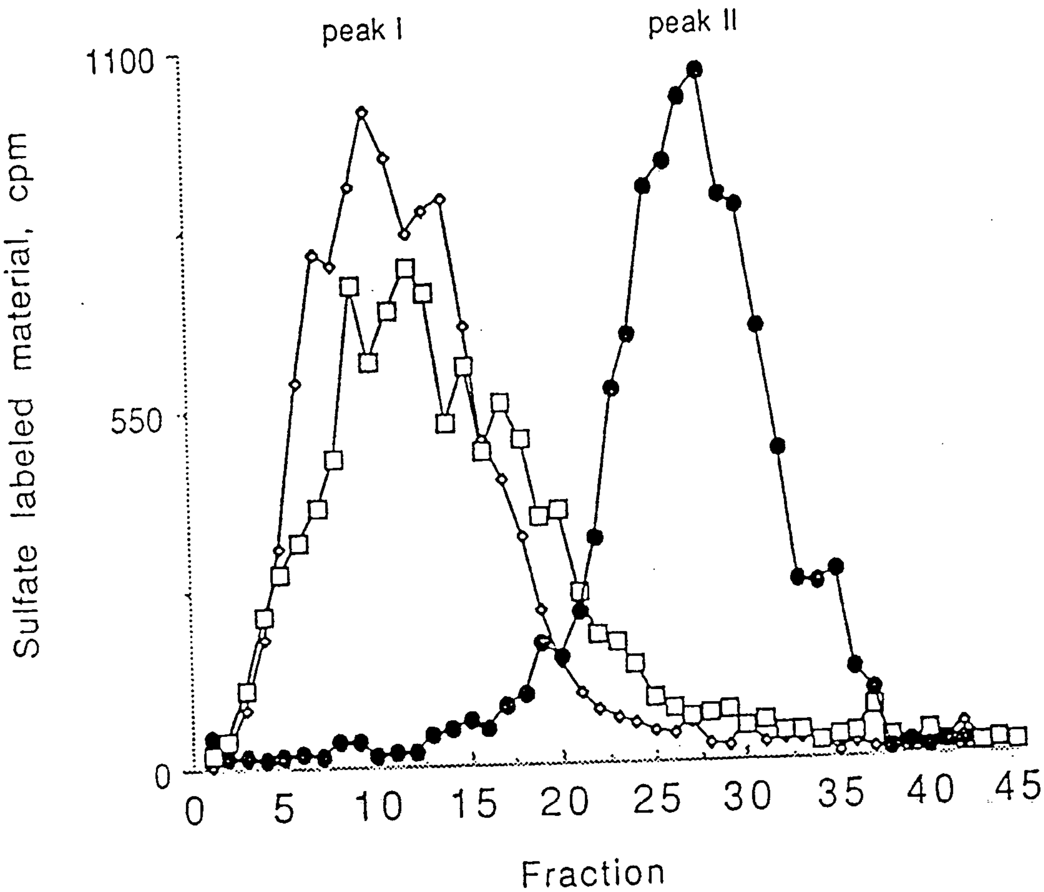
(a) hybridizing the chromosome spread with a tagged polynucleotyde probe encoding heparanase;

(b) washing the chromosome spread, thereby removing excess of non-hybridized probe; and

(c) searching for signals associated with said hybridized tagged polynucleotyde probe, wherein detected signals being indicative of a chromosome region harboring a human heparanase gene.

```
   1  CTAGAGCTTTCGACTCTCCGCTGCGCGGCAGCTGGCGGGGGGAGCAGCCAGGTGAGCCCA

  61  AGATGCTGCTGCGCTCGAAGCCTGCGCTGCCGCCGCCGCTGATGCTGCTGCTCCTGGGGC
       M  L  L  R  S  K  P  A  L  P  P  P  L  M  L  L  L  L  G  P

 121  CGCTGGGTCCCCTCTCCCCTGGCGCCCTGCCCCGACCTGCGCAAGCACAGGACGTCGTGG
       L  G  P  L  S  P  G  A  L  P  R  P  A  Q  A  Q  D  V  V  D

 181  ACCTGGACTTCTTCACCCAGGAGCCGCTGCACCTGGTGAGCCCCTCGTTCCTGTCCGTCA
       L  D  F  F  T  Q  E  P  L  H  L  V  S  P  S  F  L  S  V  T

 241  CCATTGACGCCAACCTGGCCACGGACCCGCGGTTCCTCATCCTCCTGGGTTCTCCAAAGC
       I  D  A  N  L  A  T  D  P  R  F  L  I  L  L  G  S  P  K  L

 301  TTCGTACCTTGGCCAGAGGCTTGTCTCCTGCGTACCTGAGGTTTGGTGGCACCAAGACAG
       R  T  L  A  R  G  L  S  P  A  Y  L  R  F  G  G  T  K  T  D

 361  ACTTCCTAATTTTCGATCCCAAGAAGGAATCAACCTTTGAAGAGAGAAGTTACTGGCAAT
       F  L  I  F  D  P  K  K  E  S  T  F  E  E  R  S  Y  W  Q  S

 421  CTCAAGTCAACCAGGATATTTGCAAATATGGATCCATCCCTCCTGATGTGGAGGAGAAGT
       Q  V  N  Q  D  I  C  K  Y  G  S  I  P  P  D  V  E  E  K  L

 481  TACGGTTGGAATGGCCCTACCAGGAGCAATTGCTACTCCGAGAACACTACCAGAAAAAGT
       R  L  E  W  P  Y  Q  E  Q  L  L  L  R  E  H  Y  Q  K  K  F

 541  TCAAGAACAGCACCTACTCAAGAAGCTCTGTAGATGTGCTATACACTTTTGCAAACTGCT
       K  N  S  T  Y  S  R  S  S  V  D  V  L  Y  T  F  A  N  C  S

 601  CAGGACTGGACTTGATCTTTGGCCTAAATGCGTTATTAAGAACAGCAGATTTGCAGTGGA
       G  L  D  L  I  F  G  L  N  A  L  L  R  T  A  D  L  Q  W  N

 661  ACAGTTCTAATGCTCAGTTGCTCCTGGACTACTGCTCTTCCAAGGGGTATAACATTTCTT
       S  S  N  A  Q  L  L  L  D  Y  C  S  S  K  G  Y  N  I  S  W

 721  GGGAACTAGGCAATGAACCTAACAGTTTCCTTAAGAAGGCTGATATTTTCATCAATGGGT
       E  L  G  N  E  P  N  S  F  L  K  K  A  D  I  F  I  N  G  S
                      (T)
 781  CGCAGTTAGGAGAAGATTATATTCAATTGCATAAACTTCTAAGAAAGTCCACCTTCAAAA
       Q  L  G  E  D  Y  I  Q  L  H  K  L  L  R  K  S  T  F  K  N
                     (F)
 841  ATGCAAAACTCTATGGTCCTGATGTTGGTCAGCCTCGAAGAAAGACGGCTAAGATGCTGA
       A  K  L  Y  G  P  D  V  G  Q  P  R  R  K  T  A  K  M  L  K

 901  AGAGCTTCCTGAAGGCTGGTGGAGAAGTGATTGATTCAGTTACATGGCATCACTACTATT
       S  F  L  K  A  G  G  E  V  I  D  S  V  T  W  H  H  Y  Y  L

 961  TGAATGGACGGACTGCTACCAGGGAAGATTTTCTAAACCCTGATGTATTGGACATTTTTA
       N  G  R  T  A  T  R  E  D  F  L  N  P  D  V  L  D  I  F  I

1021  TTTCATCTGTGCAAAAAGTTTTCCAGGTGGTTGAGAGCACCAGGCCTGGCAAGAAGGTCT
       S  S  V  Q  K  V  F  Q  V  V  E  S  T  R  P  G  K  K  V  W

1081  GGTTAGGAGAAACAAGCTCTGCATATGGAGGCGGAGCGCCCTTGCTATCCGACACCTTTG
       L  G  E  T  S  S  A  Y  G  G  G  A  P  L  L  S  D  T  F  A

1141  CAGCTGGCTTTATGTGGCTGGATAAATTGGGCCTGTCAGCCCGAATGGGAATAGAAGTGG
       A  G  F  M  W  L  D  K  L  G  L  S  A  R  M  G  I  E  V  V

1201  TGATGAGGCAAGTATTCTTTGGAGCAGGAAACTACCATTTAGTGGATGAAAACTTCGATC
       M  R  Q  V  F  F  G  A  G  N  Y  H  L  V  D  E  N  F  D  P

1261  CTTTACCTGATTATTGGCTATCTCTTCTGTTCAAGAAATTGGTGGGCACCAAGGTGTTAA
       L  P  D  Y  W  L  S  L  L  F  K  K  L  V  G  T  K  V  L  M

1321  TGGCAAGCGTGCAAGGTTCAAAGAGAAGGAAGCTTCGAGTATACCTTCATTGCACAAACA
       A  S  V  Q  G  S  K  R  R  K  L  R  V  Y  L  H  C  T  N  T

1381  CTGACAATCCAAGGTATAAAGAAGGAGATTTAACTCTGTATGCCATAAACCTCCATAACG
       D  N  P  R  Y  K  E  G  D  L  T  L  Y  A  I  N  L  H  N  V

1441  TCACCAAGTACTTGCGGTTACCCTATCCTTTTTCTAACAAGCAAGTGGATAAATACCTTC
       T  K  Y  L  R  L  P  Y  P  F  S  N  K  Q  V  D  K  Y  L  L

1501  TAAGACCTTTGGGACCTCATGGATTACTTTCCAAATCTGTCCAACTCAATGGTCTAACTC
       R  P  L  G  P  H  G  L  L  S  K  S  V  Q  L  N  G  L  T  L

1561  TAAAGATGGTGGATGATCAAACCTTGCCACCTTTAATGGAAAAACCTCTCCGGCCAGGAA
       K  M  V  D  D  Q  T  L  P  P  L  M  E  K  P  L  R  P  G  S

1621  GTTCACTGGGCTTGCCAGCTTTCTCATATAGTTTTTTTGTGATAAGAAATGCCAAAGTTG
       S  L  G  L  P  A  F  S  Y  S  F  F  V  I  R  N  A  K  V  A
```

FIG. 2



peak I
peak II
1100
550
0
Sulfate labeled material, cpm
0
5
10
15
20
25
30
35
40
45
Fraction

FIG. 3A

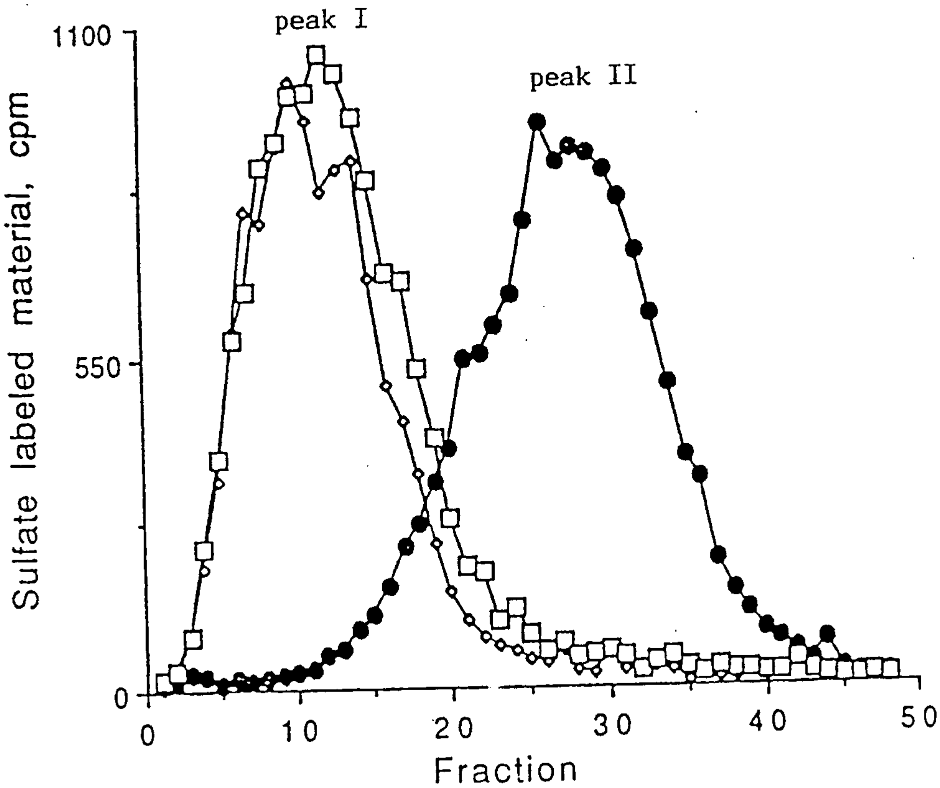

peak I
peak II
1100
550
0
0
10
20
30
40
50
Sulfate labeled material, cpm
Fraction

FIG. 3B

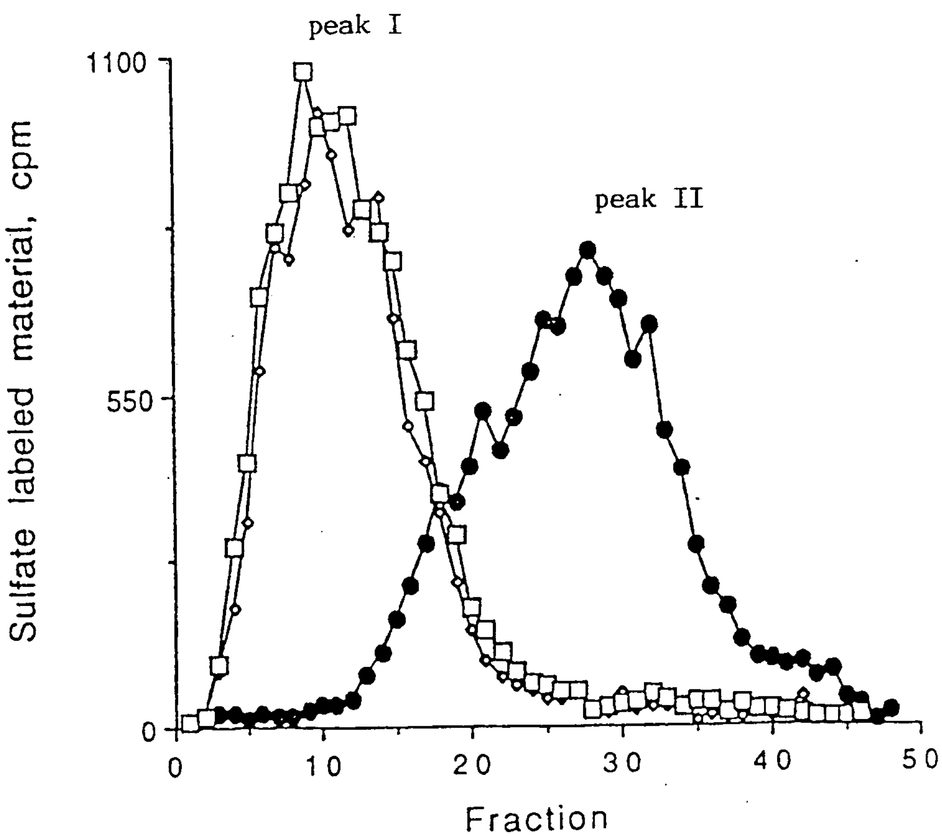

peak I
peak II
1100
550
0
0
10
20
30
40
50
Sulfate labeled material, cpm
Fraction

FIG. 4

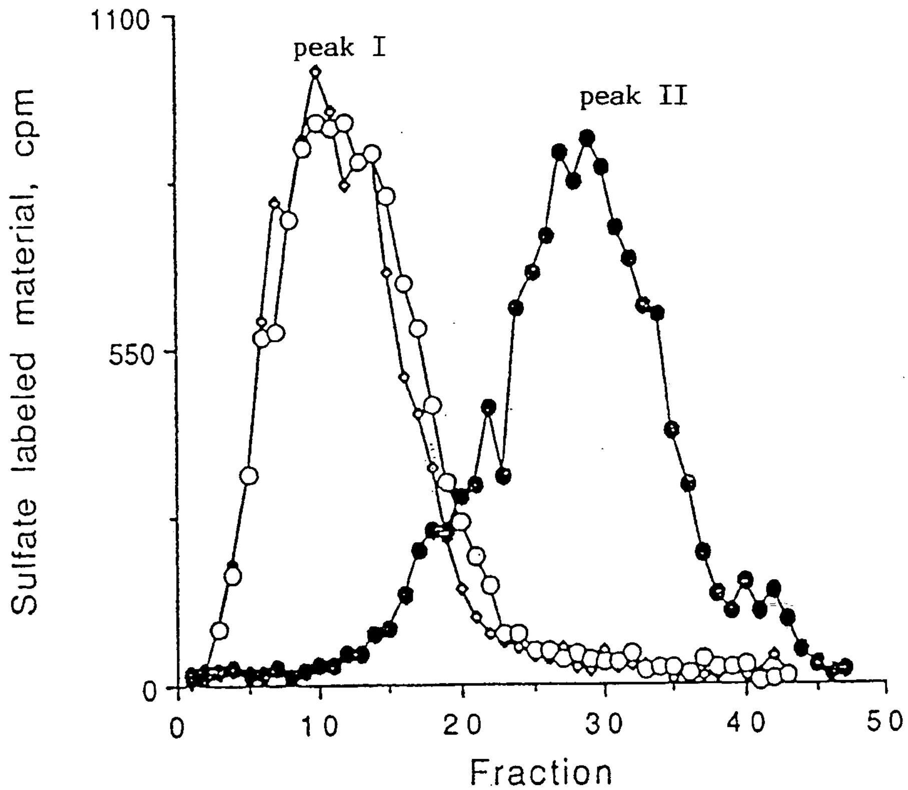
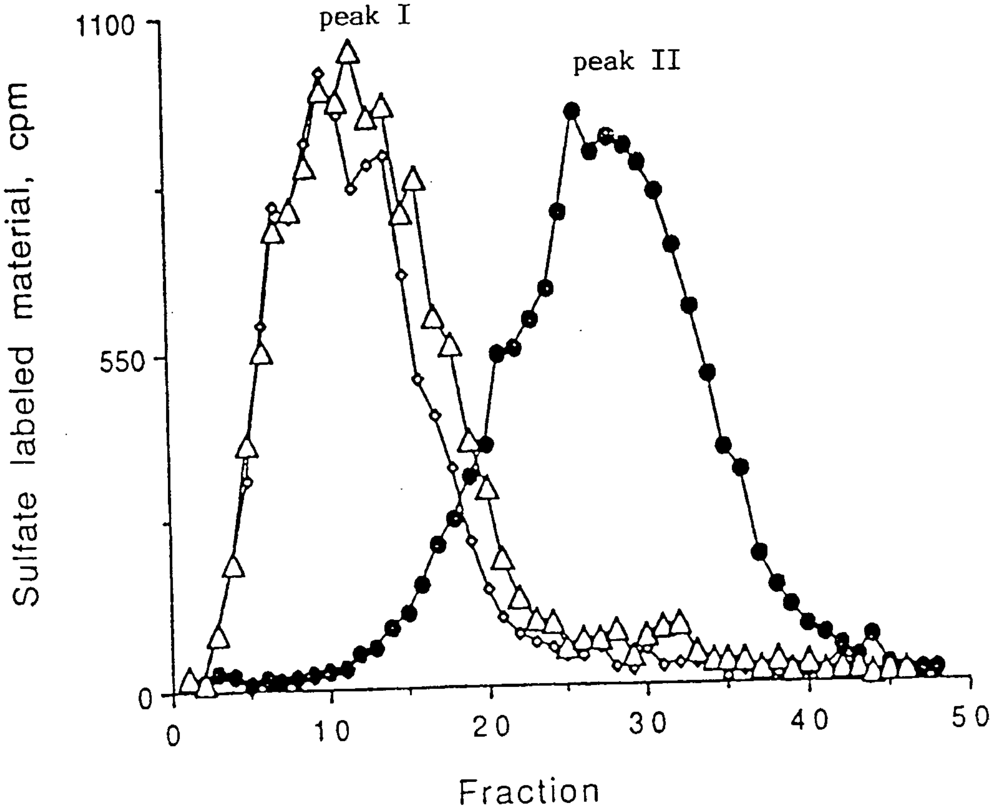
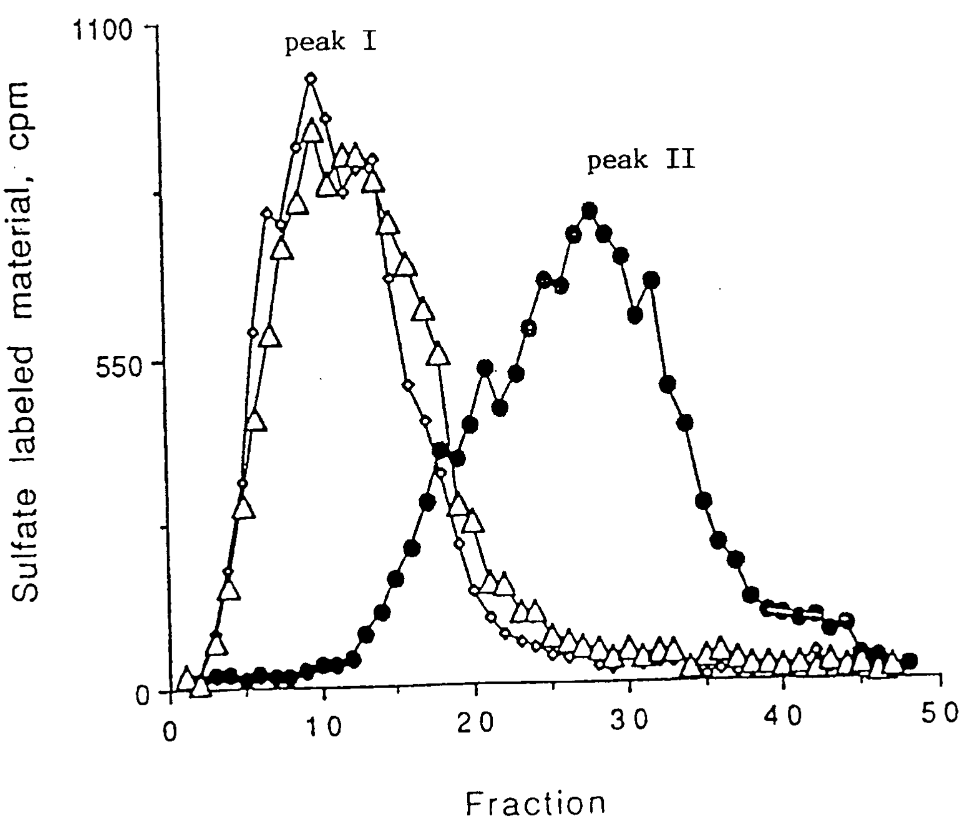

peak I
peak II
Sulfate labeled material, cpm
1100
550
0
0
10
20
30
40
50
Fraction

FIG. 5A



peak I
peak II
1100
550
0
Sulfate labeled material, cpm
0
10
20
30
40
50
Fraction

FIG. 5B



peak I
peak II
1100
550
0
Sulfate labeled material, cpm
0
10
20
30
40
50
Fraction

FIG. 6A

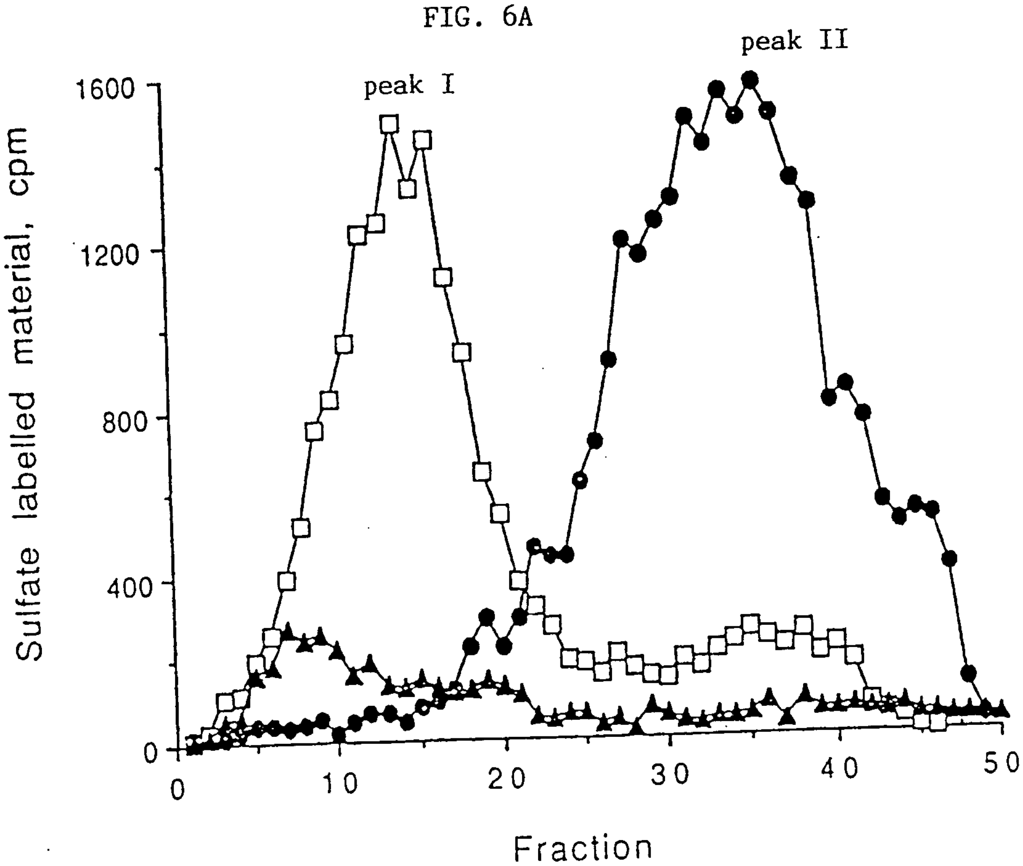

peak I
peak II
1600
1200
800
400
0
Sulfate labelled material, cpm
0
10
20
30
40
50
Fraction

FIG. 6B

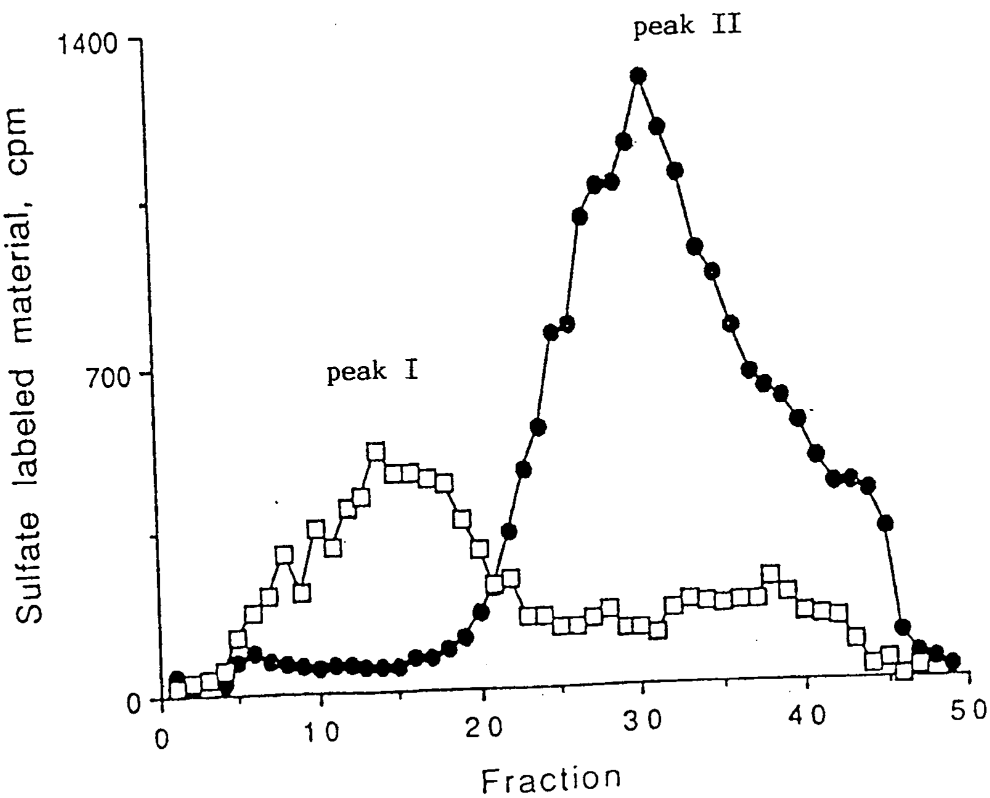

peak II
peak I
1400
700
0
Sulfate labeled material, cpm
0
10
20
30
40
50
Fraction

FIG. 7A

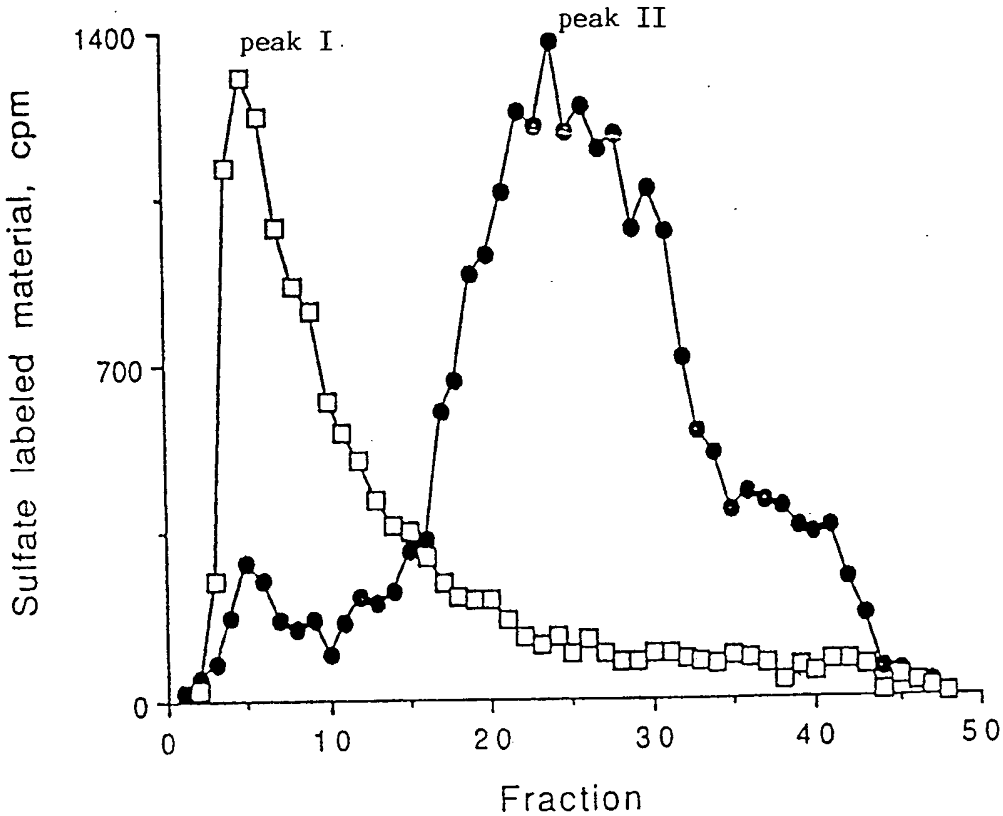

peak I
peak II
1400
700
0
Sulfate labeled material, cpm
0
10
20
30
40
50
Fraction

FIG. 7B

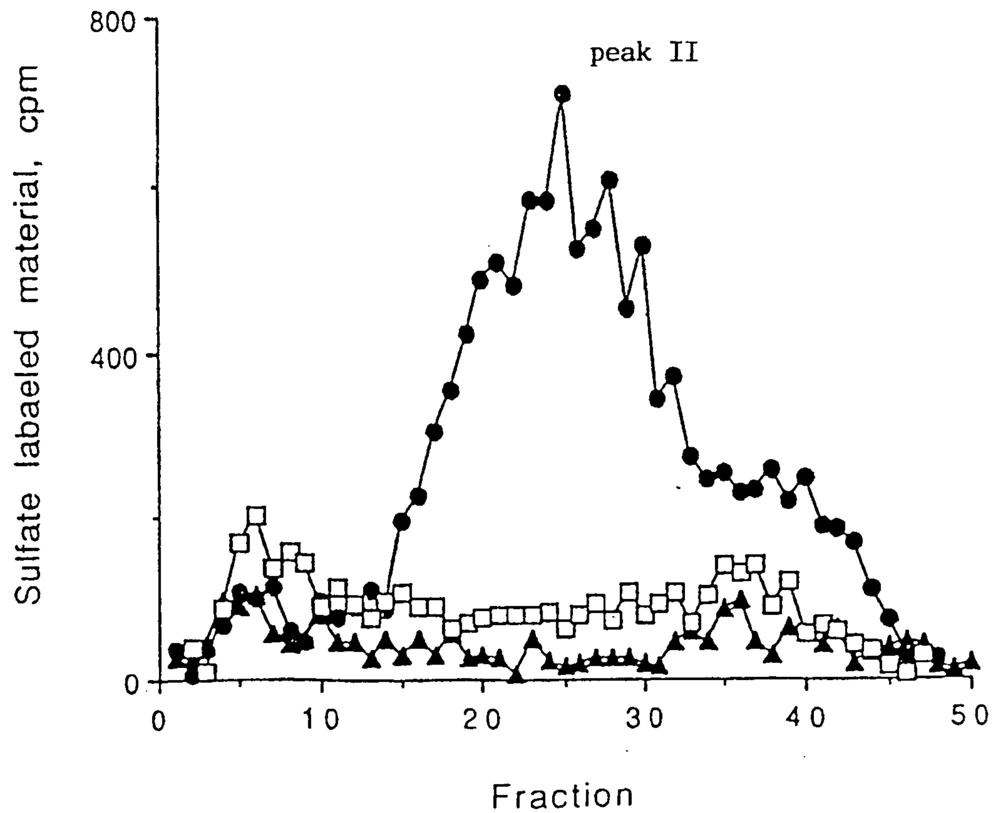

peak II
800
400
0
Sulfate labaeled material, cpm
0
10
20
30
40
50
Fraction

FIG. 8A

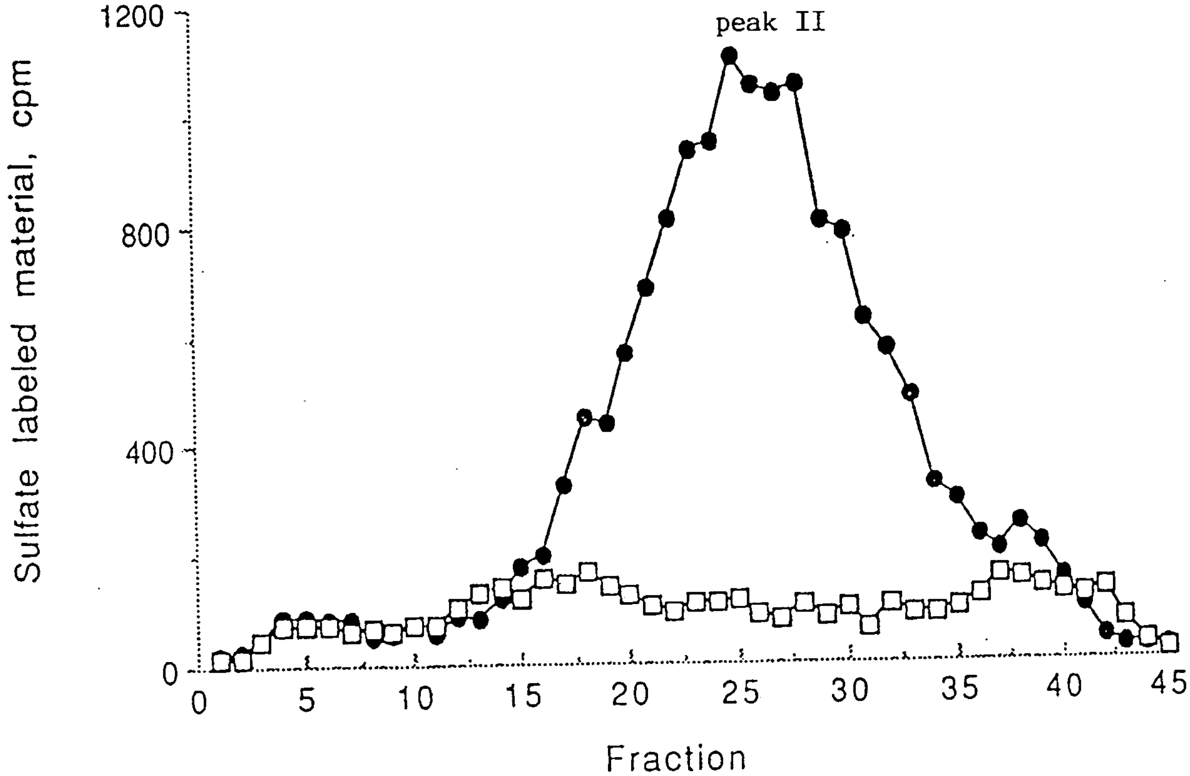

peak II
1200
800
400
0
Sulfate labeled material, cpm
0
5
10
15
20
25
30
35
40
45
Fraction

FIG. 8B

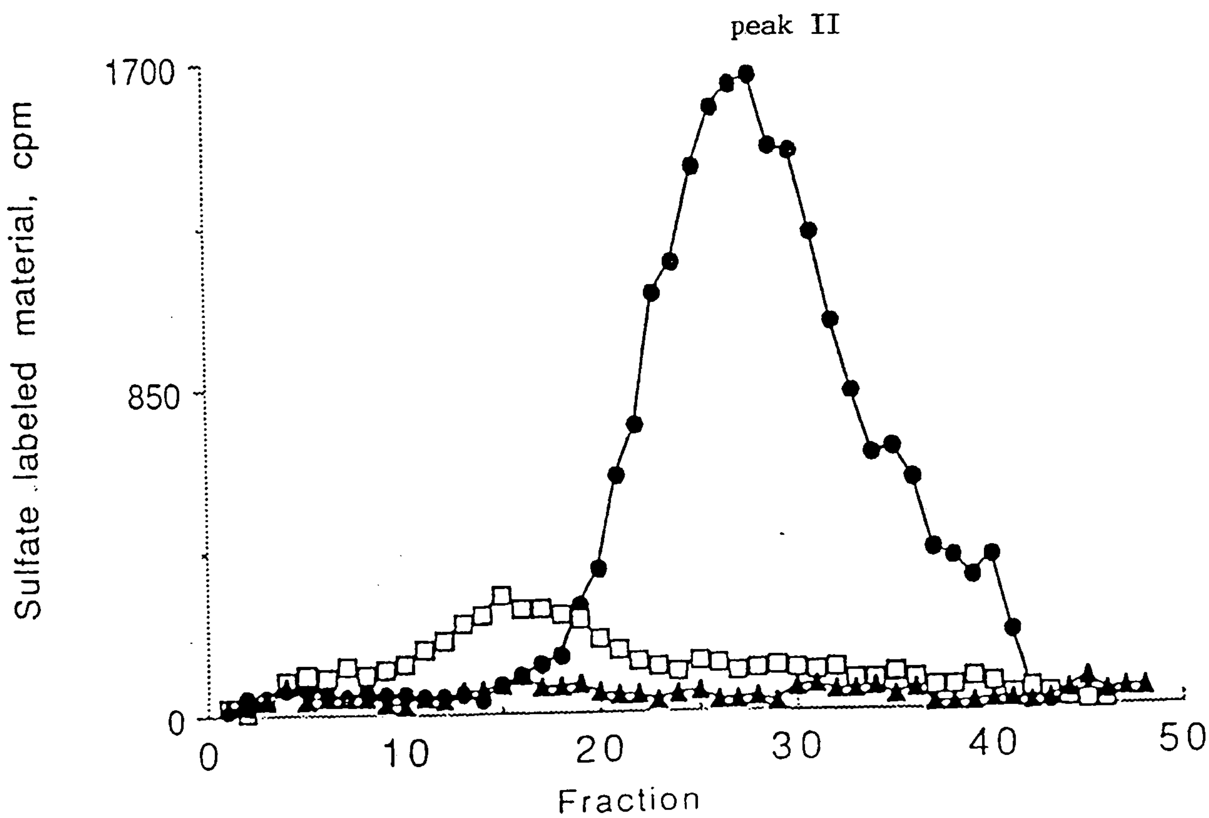

peak II
1700
850
0
Sulfate labeled material, cpm
0
10
20
30
40
50
Fraction

FIG. 9A

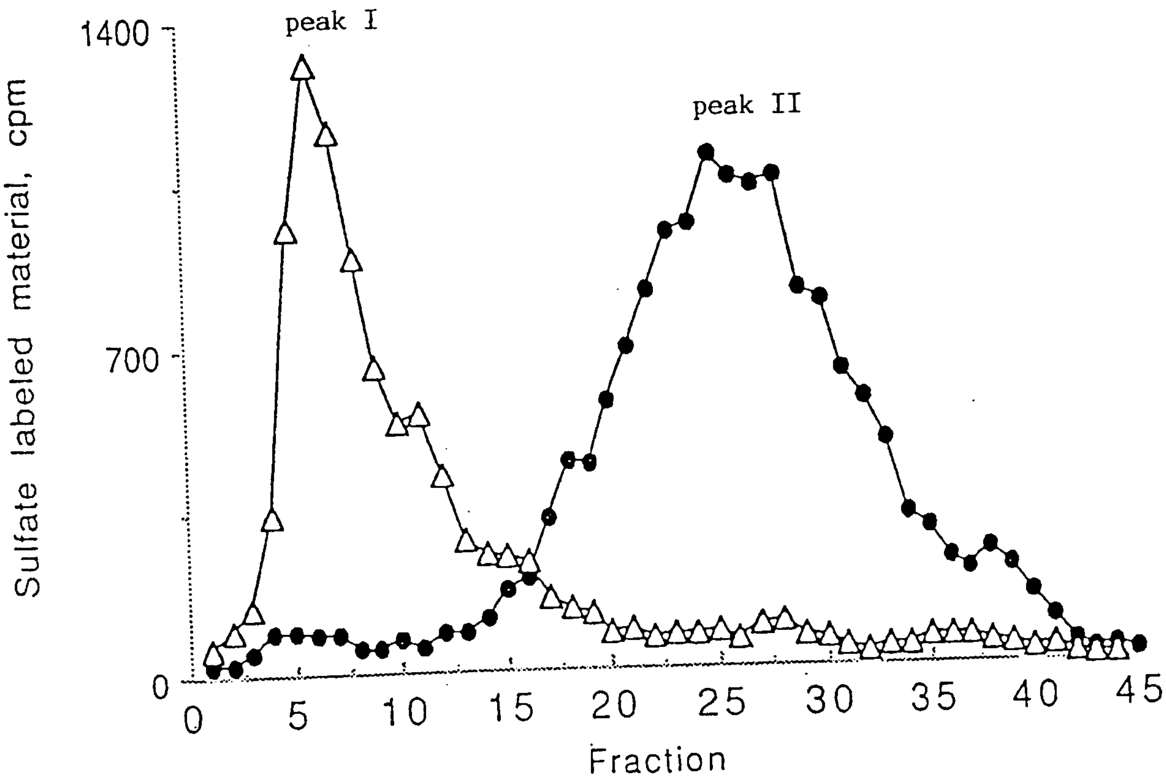

peak I
peak II
1400
700
0
Sulfate labeled material, cpm
0 5 10 15 20 25 30 35 40 45
Fraction

FIG. 9B

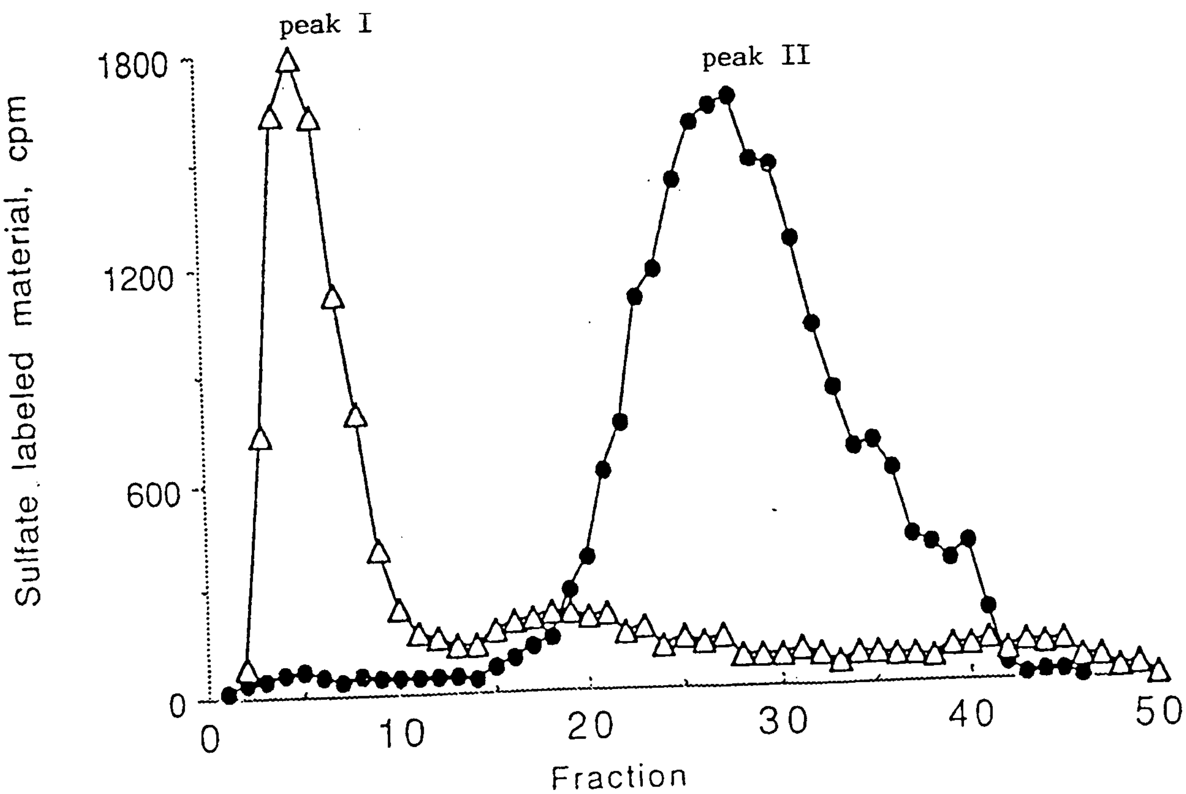

peak I
peak II
1800
1200
600
0
Sulfate labeled material, cpm
0 10 20 30 40 50
Fraction

FIG. 10A

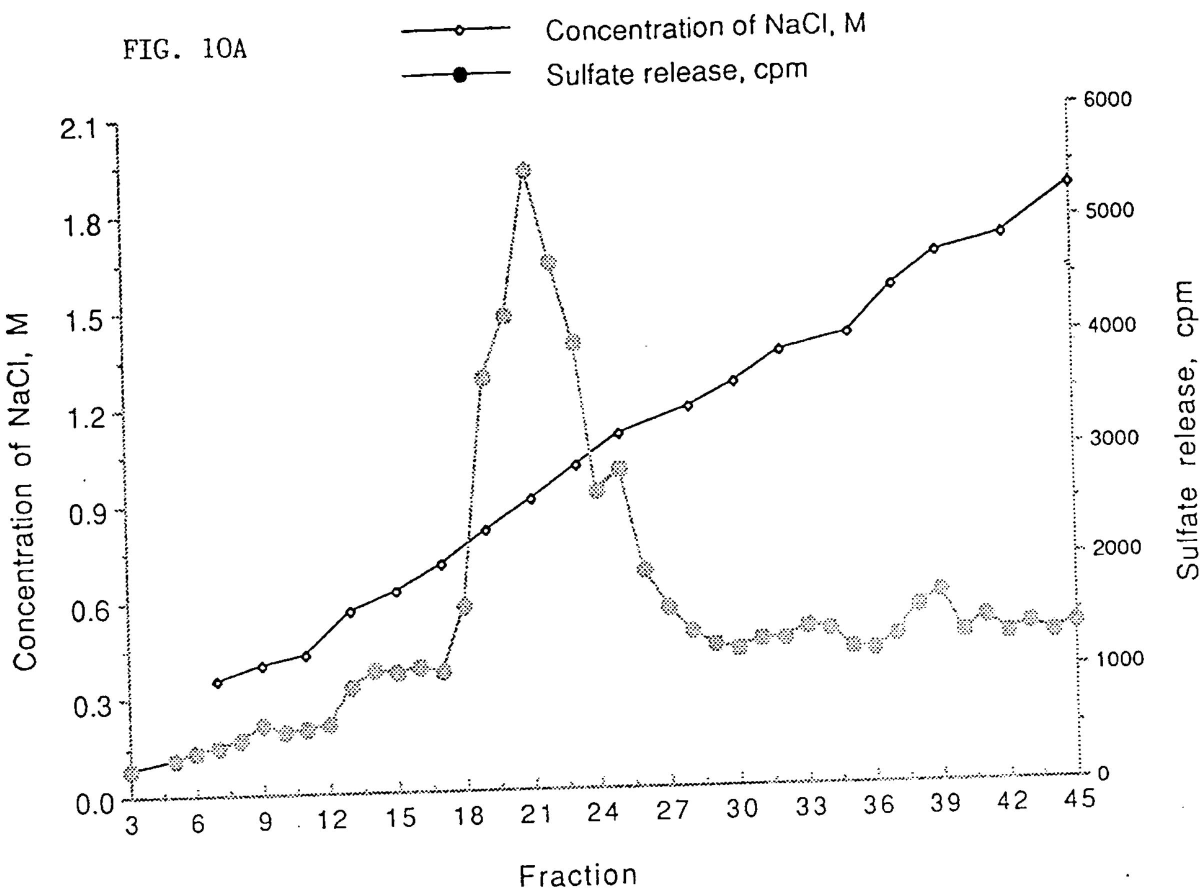


FIG. 10B

Heparin-avidogel

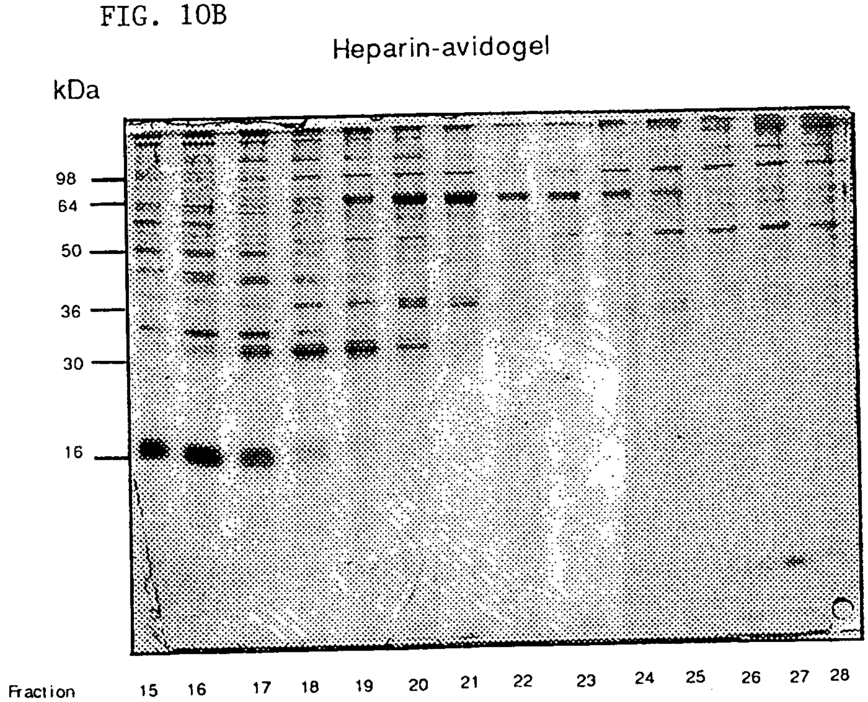


FIG. 11A

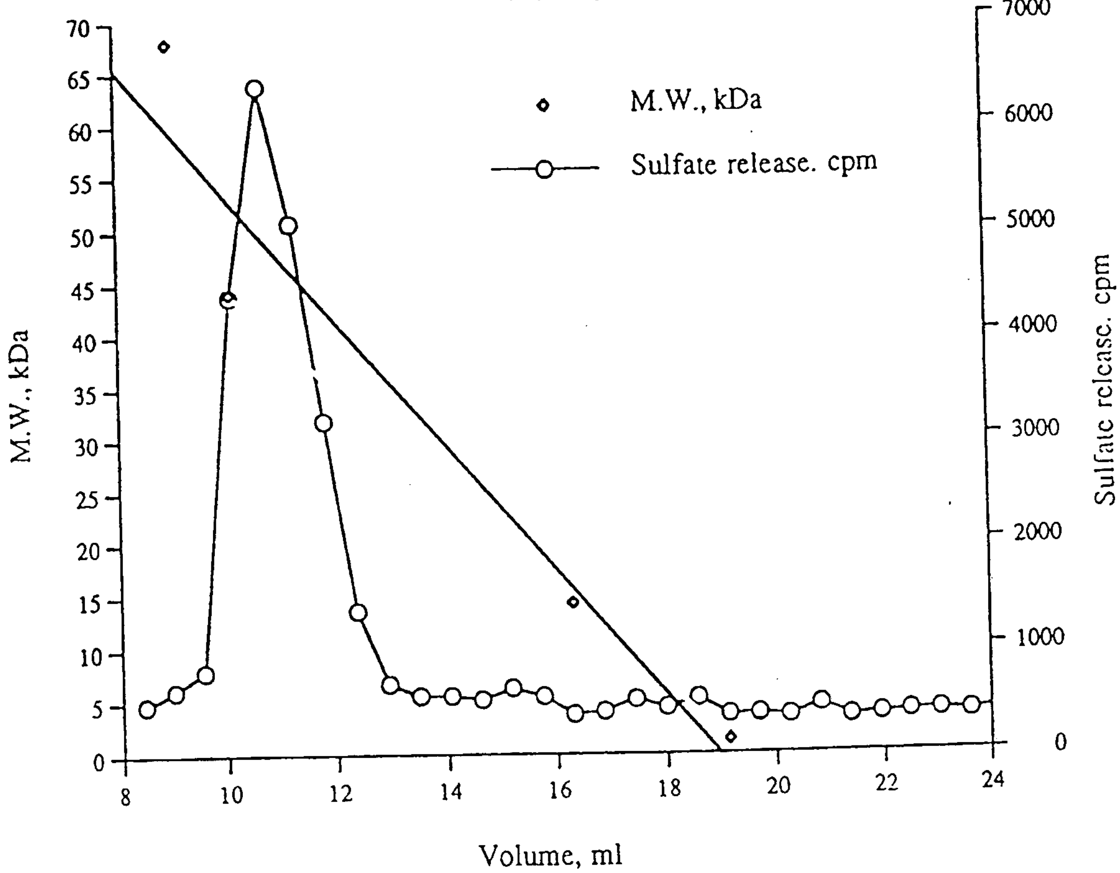


FIG. 11B

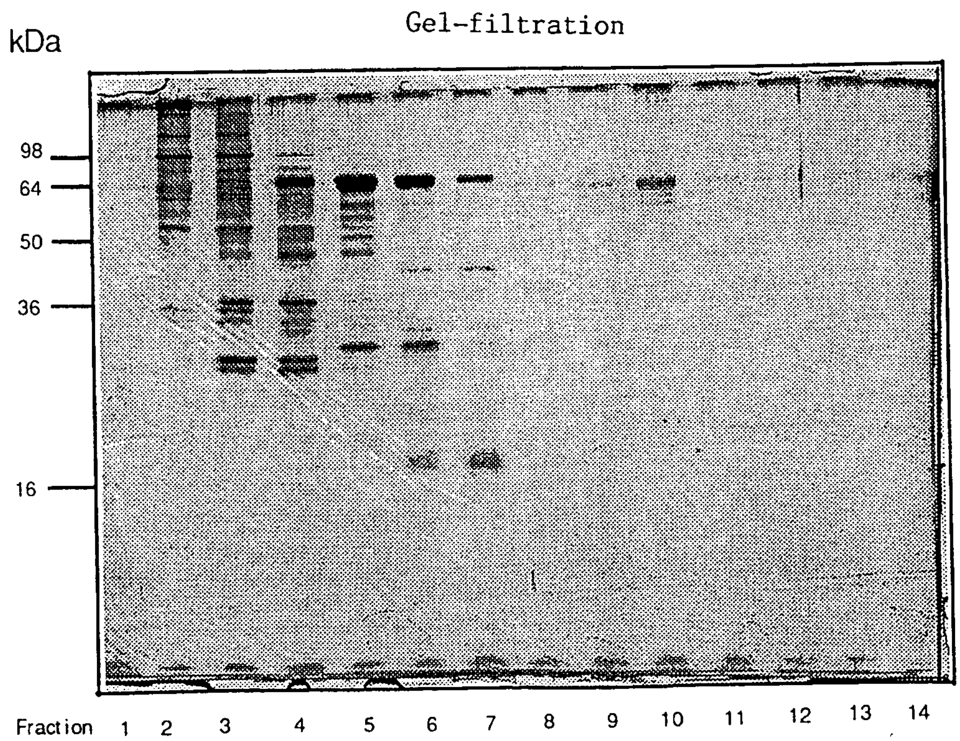


FIG. 12A

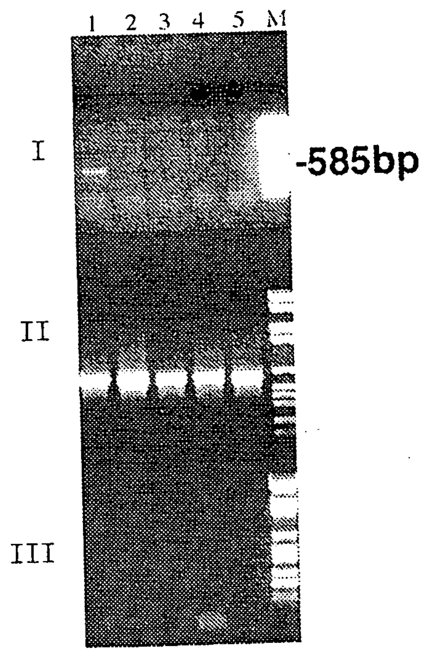


FIG. 12B

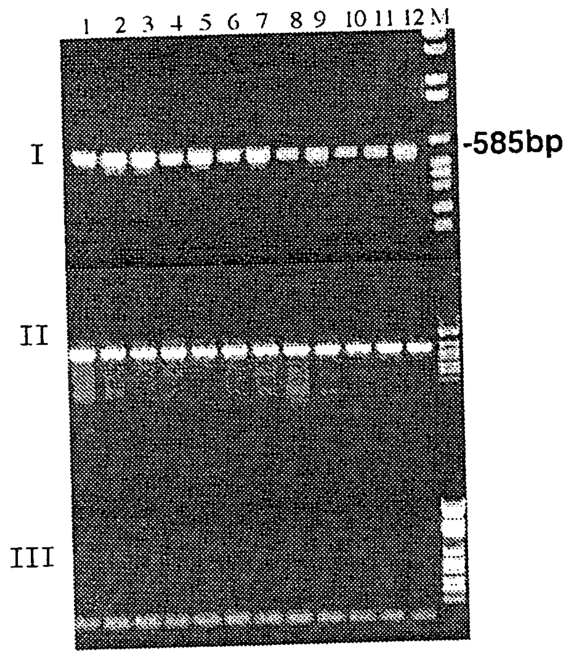


FIG. 12C

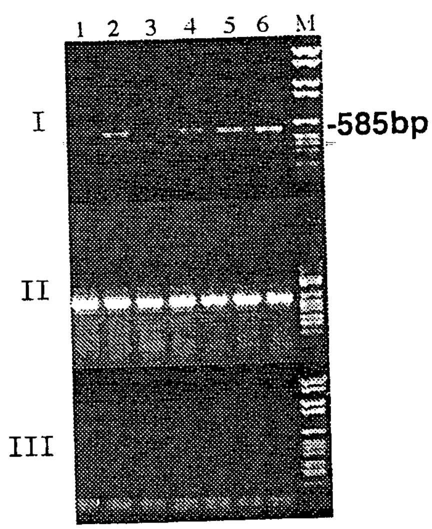


FIG. 12D

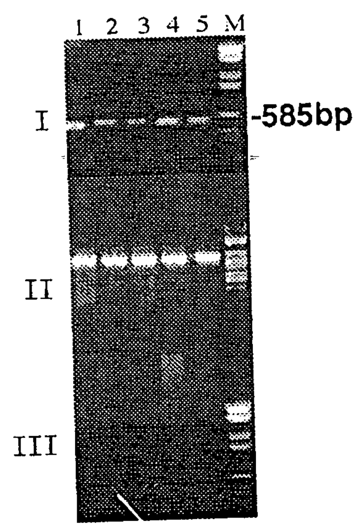


FIG. 12E

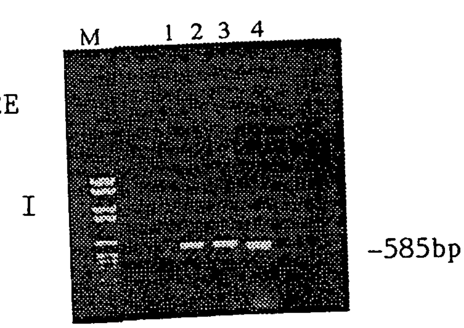


Fig. 13

```
mouse  CTGGCAAGAAGGTCTGGTTGGGAGAGACGAGCTCAGCTTACGGTGGCGGT 50
human  CTGGCAAGAAGGTCTGGTTAGGAGAAACAAGCTCTGCATATGGAGGCGGA 1115

mouse  GCACCCTTGCTGTCCAACACCTTTGCAGCTGGCTTTATGTGGCTGGATAA 100
human  GCGCCCTTGCTATCCGACACCTTTGCAGCTGGCTTTATGTGGCTGGATAA 1165

mouse  ATTGGGCCTGTCAGCCCAGATGGGCATAGAAGTCGTGATGAGGCAGGTGT 150
human  ATTGGGCCTGTCAGCCCGAATGGGAATAGAAGTGGTGATGAGGCAAGTAT 1215

mouse  TCTTCGGAGCAGGCAACTACCACTTAGTGGATGAAAACTTTGAGCCTTTA 200
human  TCTTTGGAGCAGGAAACTACCATTTAGTGGATGAAAACTTCGATCCTTTA 1265

mouse  CCTGATTACTGGCTCTCTCTTCTGTTCAAGAAACTGGTAGGTCCCAGGGT 250
human  CCTGATTATTGGCTATCTCTTCTGTTCAAGAAATTGGTGGGCACCAAGGT 1315

mouse  GTTACTGTCAAGAGTGAAAGGCCCAGACAGGAGCAAACTCCGAGTGTATC 300
human  GTTAATGGCAAGCGTGCAAGGTTCAAAGAGAAGGAAGCTTCGAGTATACC 1365

mouse  TCCACTGCACTAACGTCTATCACCCACGATATCAGGAAGGAGATCTAACT 350
human  TTCATTGCACAAACACTGACAATCCAAGGTATAAAGAAGGAGATTTAACT 1415

mouse  CTGTATGTCCTGAACCTCCATAATGTCACCAAGCACTTGAAGGTACCGCC 400
human  CTGTATGCCATAAACCTCCATAACGTCACCAAGTACTTGCGGTTACCCTA 1465

mouse  TCCGTTGTTCAGGAAACCAGTGGATACGTACCTTCTGAAGCCTTCGGGGC 450
human  TCCTTTTTCTAACAAGCAAGTGGATAAATACCTTCTAAGACCTTTGGGAC 1515

mouse  CGGATGGATTACTTTCCAAATCTGTCCAACTGAACGGTCAAATTCTGAAG 500
human  CTCATGGATTACTTTCCAAATCTGTCCAACTCAATGGTCTAACTCTAAAG 1565

mouse  ATGGTGGATGAGCAGACCCTGCCAGCTTTGACAGAAAAACCTCTCCCCGC 550
human  ATGGTGGATGATCAAACCTTGCCACCTTTAATGGAAAAACCTCTCCGGCC 1615

mouse  AGGAAGTGCACTAAGCCTGCCTGCCTTTTCCTATGGTTTTTTTGTCATAA 600
human  AGGAAGTTCACTGGGCTTGCCAGCTTTCTCATATAGTTTTTTTGTGATAA 1665

mouse  GAAATGCCAAAATCGCTGCTTGTATATGAAAATAAAA 637
human  GAAATGCCAAAGTTGCTGCTTGCATCTGAAAATAAAA 1702
```

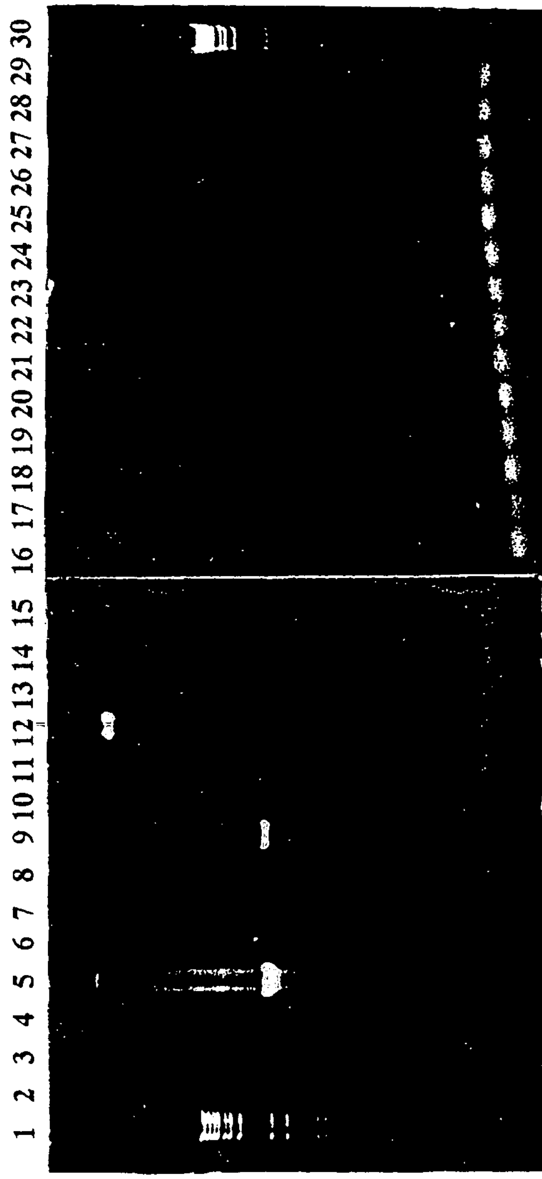

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

FIG. 14

SEQUENCE LISTING

(1) GENERAL INFORMATION:
(i) APPLICANT: Iris Pecker, Israel Vlodavsky and Elena Feinstein
(ii) TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY AND EXPRESSION OF SAME IN TRANSDUCED CELLS
(iii) NUMBER OF SEQUENCES: 23
(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Mark M. Friedman c/o Robert Sheinbein
(B) STREET: 2940 Birchtree lane
(C) CITY: Silver Spring
(D) STATE: Maryland
(E) COUNTRY: United States of America
(F) ZIP: 20906
(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
(B) COMPUTER: Twinhead* Slimnote-890TX
(C) OPERATING SYSTEM: MS DOS version 6.2, Windows version 3.11
(D) SOFTWARE: Word for Windows version 2.0 converted to an ASCI file
(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:
(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/922,170
(B) FILING DATE: 2 SEP 1997
(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Friedmam, Mark M.
(B) REGISTRATION NUMBER: 33,883
(C) REFERENCE/DOCKET NUMBER: 910/1
(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 972-3-5625553
(B) TELEFAX: 972-3-5625554
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
CCATCCTAAT ACGACTCACT ATAGGGC 27

(2) INFORMATION FOR SEQ ID NO:2:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
GTAGTGATGC CATGTAACTG AATC 24

(2) INFORMATION FOR SEQ ID NO:3:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
ACTCACTATA GGGCTCGAGC GGC 23

(2) INFORMATION FOR SEQ ID NO:4:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
GCATCTTAGC CGTCTTTCTT CG 22

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:
TTTTTTTTTT TTTTT 15

(2) INFORMATION FOR SEQ ID NO:6:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:
TTCGATCCCA AGAAGGAATC AAC 23

(2) INFORMATION FOR SEQ ID NO:7:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:
GTAGTGATGC CATGTAACTG AATC 24

(2) INFORMATION FOR SEQ ID NO:8:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:
Tyr Gly Pro Asp Val Gly Gln Pro Arg
5 9

(2) INFORMATION FOR SEQ ID NO:9:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1721
(B) TYPE: nucleic acid
(G) STRANDEDNESS: double
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
CTAGAGCTTT CGACTCTCCG CTGCGCGGCA GCTGGCGGGG GGAGCAGCCA GGTGAGCCCA 60
AGATGCTGCT GCGCTCGAAG CCTGCGCTGC CGCCGCCGCT GATGCTGCTG CTCCTGGGGC 120
CGCTGGGTCC CCTCTCCCCT GGCGCCCTGC CCCGACCTGC GCAAGCACAG GACGTCGTGG 180
ACCTGGACTT cTTCACCCAG GAGCCGCTGC ACCTGGTGAG CCCCTCGTTC CTGTCCGTCA 240
CCATTGACGC CAACCTGGCC ACGGACCCGC GGTTCCTCAT CCTCCTGGGT TCTCCAAAGC 300
TTCGTACCTT GGCCAGAGGC TTGTCTCCTG CGTACCTGAG GTTTGGTGGC ACCAAGACAG 360
ACTTCCTAAT TTTCGATCCC AAGAAGGAAT CAACCTTTGA AGAGAGAAGT TACTGGCAAT 420
CTCAAGTCAA CCAGGATATT TGCAAATATG GATCCATCCC TCCTGATGTG GAGGAGAAGT 480
TACGGTTGGA ATGGCCCTAC CAGGAGCAAT TGCTACTCCG AGAACACTAC CAGAAAAAGT 540
TCAAGAACAG CACCTACTCA AGAAGCTCTG TAGATGTGCT ATACACTTTT GCAAACTGCT 600
CAGGACTGGA CTTGATCTTT GGCCTAAATG CGTTATTAAG AACAGCAGAT TTGCAGTGGA 660
ACAGTTCTAA TGCTCAGTTG CTCCTGGACT ACTGCTCTTC CAAGGGGTAT AACATTTCTT 720
GGGAACTAGG CAATGAACCT AACAGTTTCC TTAAGAAGGC TGATATTTTC ATCAATGGGT 780
CGCAGTTAGG AGAAGATTAT ATTCAATTGC ATAAACTTCT AAGAAAGTCC ACCTTCAAAA 840
ATGCAAAACT CTATGGTCCT GATGTTGGTC AGCCTCGAAG AAAGACGGCT AAGATGCTGA 900
AGAGCTTCCT GAAGGCTGGT GGAGAAGTGA TTGATTCAGT TACATGGCAT CACTACTATT 960
TGAATGGACG GACTGCTACC AGGGAAGATT TTCTAAACCC TGATGTATTG GACATTTTTA 1020
TTTCATCTGT GCAAAAAGTT TTCCAGGTGG TTGAGAGCAC CAGGCCTGGC AAGAAGGTCT 1080
GGTTAGGAGA AACAAGCTCT GCATATGGAG GCGGAGCGCC CTTGCTATCC GACACCTTTG 1140
CAGCTGGCTT TATGTGGCTG GATAAATTGG GCCTGTCAGC CCGAATGGGA ATAGAAGTGG 1200
TGATGAGGCA AGTATTCTTT GGAGCAGGAA ACTACCATTT AGTGGATGAA AACTTCGATC 1260
CTTTACCTGA TTATTGGCTA TCTCTTCTGT TCAAGAAATT GGTGGGCACC AAGGTGTTAA 1320
TGGCAAGCGT GCAAGGTTCA AAGAGAAGGA AGCTTCGAGT ATACCTTCAT TGCACAAACA 1380
CTGACAATCC AAGGTATAAA GAAGGAGATT TAACTCTGTA TGCCATAAAC CTCCATAACG 1440
TCACCAAGTA CTTGCGGTTA CCCTATCCTT TTTCTAACAA GCAAGTGGAT AAATACCTTC 1500
TAAGACCTTT GGGACCTCAT GGATTACTTT CCAAATCTGT CCAACTCAAT GGTCTAACTC 1560
TAAAGATGGT GGATGATCAA ACCTTGCCAC CTTTAATGGA AAAACCTCTC CGGCCAGGAA 1620
GTTCACTGGG CTTGCCAGCT TTCTCATATA GTTTTTTTGT GATAAGAAAT GCCAAAGTTG 1680
CTGCTTGCAT CTGAAAATAA AATATACTAG TCCTGACACT G                     1721
```

(2) INFORMATION FOR SEQ ID NO:10:
(i) SEQUENCE CHARACTERISTICS:

III

(A) LENGTH: 543
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro Pro Leu Met Leu Leu
                5                   10                  15

Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro Gly Ala Leu Pro Arg Pro
            20                  25                  30

Ala Gln Ala Gln Asp Val Val Asp Leu Asp Phe Phe Thr Gln Glu Pro
        35                  40                  45

Leu His Leu Val Ser Pro Ser Phe Leu Ser Val Thr Ile Asp Ala Asn
    50                  55                  60

Leu Ala Thr Asp Pro Arg Phe Leu Ile Leu Leu Gly Ser Pro Lys Leu
65                  70                  75                  80

Arg Thr Leu Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly
                85                  90                  95

Thr Lys Thr Asp Phe Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe
            100                 105                 110

Glu Glu Arg Ser Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys
        115                 120                 125

Tyr Gly Ser Ile Pro Pro Asp Val Glu Glu Lys Leu Arg Leu Glu Trp
    130                 135                 140

Pro Tyr Gln Glu Gln Leu Leu Leu Arg Glu His Tyr Gln Lys Lys Phe
145                 150                 155                 160

Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe
                165                 170                 175

Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu Leu
            180                 185                 190

Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu Leu Leu
        195                 200                 205

Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu Leu Gly Asn
    210                 215                 220

Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe Ile Asn Gly Ser
225                 230                 235                 240

Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys Leu Leu Arg Lys Ser
                245                 250                 255

Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro Asp Val Gly Gln Pro Arg
            260                 265                 270

Arg Lys Thr Ala Lys Met Leu Lys Ser Phe Leu Lys Ala Gly Gly Glu
        275                 280                 285

Val Ile Asp Ser Val Thr Trp His His Tyr Tyr Leu Asn Gly Arg Thr
    290                 295                 300

Ala Thr Arg Glu Asp Phe Leu Asn Pro Asp Val Leu Asp Ile Phe Ile
305                 310                 315                 320

Ser Ser Val Gln Lys Val Phe Gln Val Val Glu Ser Thr Arg Pro Gly
                325                 330                 335

Lys Lys Val Trp Leu Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala
            340                 345                 350

Pro Leu Leu Ser Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys
        355                 360                 365

Leu Gly Leu Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val
```

IV

370 375 380

Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro
385 390 395 400

Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr
405 410 415

Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu Arg
420 425 430

Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys Glu Gly
435 440 445

Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr Lys Tyr Leu
450 455 460

Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp Lys Tyr Leu Leu
465 470 475 480

Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys Ser Val Gln Leu Asn
485 490 495

Gly Leu Thr Leu Lys Met Val Asp Asp Gln Thr Leu Pro Pro Leu Met
500 505 510

Glu Lys Pro Leu Arg Pro Gly Ser Ser Leu Gly Leu Pro Ala Phe Ser
515 520 525

Tyr Ser Phe Phe Val Ile Arg Asn Ala Lys Val Ala Ala Cys Ile
530 535 540 543

(2) INFORMATION FOR SEQ ID NO:11:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1718
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CT AGA GCT TTC GAC 14

TCT CCG CTG CGC GGC AGC TGG CGG GGG GAG CAG CCA GGT GAG CCC AAG 62

ATG CTG CTG CGC TCG AAG CCT GCG CTG CCG CCG CCG CTG ATG CTG CTG 110
Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro Pro Leu Met Leu Leu
5 10 15

CTC CTG GGG CCG CTG GGT CCC CTC TCC CCT GGC GCC CTG CCC CGA CCT 158
Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro Gly Ala Leu Pro Arg Pro
20 25 30

GCG CAA GCA CAG GAC GTC GTG GAC CTG GAC TTC TTC ACC CAG GAG CCG 206
Ala Gln Ala Gln Asp Val Val Asp Leu Asp Phe Phe Thr Gln Glu Pro
35 40 45

CTG CAC CTG GTG AGC CCC TCG TTC CTG TCC GTC ACC ATT GAC GCC AAC 254
Leu His Leu Val Ser Pro Ser Phe Leu Ser Val Thr Ile Asp Ala Asn
50 55 60

CTG GCC ACG GAC CCG CGG TTC CTC ATC CTC CTG GGT TCT CCA AAG CTT 302
Leu Ala Thr Asp Pro Arg Phe Leu Ile Leu Leu Gly Ser Pro Lys Leu
65 70 75 80

CGT ACC TTG GCC AGA GGC TTG TCT CCT GCG TAC CTG AGG TTT GGT GGC 350
Arg Thr Leu Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly
85 90 95

ACC AAG ACA GAC TTC CTA ATT TTC GAT CCC AAG AAG GAA TCA ACC TTT 398
Thr Lys Thr Asp Phe Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe
100 105 110

GAA GAG AGA AGT TAC TGG CAA TCT CAA GTC AAC CAG GAT ATT TGC AAA 446
Glu Glu Arg Ser Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys
115 120 125

V

TAT GGA TCC ATC CCT CCT GAT GTG GAG GAG AAG TTA CGG TTG GAA TGG 494
Tyr Gly Ser Ile Pro Pro Asp Val Glu Glu Lys Leu Arg Leu Glu Trp
130 135 140

CCC TAC CAG GAG CAA TTG CTA CTC CGA GAA CAC TAC CAG AAA AAG TTC 542
Pro Tyr Gln Glu Gln Leu Leu Leu Arg Glu His Tyr Gln Lys Lys Phe
145 150 155 160

AAG AAC AGC ACC TAC TCA AGA AGC TCT GTA GAT GTG CTA TAC ACT TTT 590
Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe
165 170 175

GCA AAC TGC TCA GGA CTG GAC TTG ATC TTT GGC CTA AAT GCG TTA TTA 638
Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu Leu
180 185 190

AGA ACA GCA GAT TTG CAG TGG AAC AGT TCT AAT GCT CAG TTG CTC CTG 686
Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu Leu Leu
195 200 205

GAC TAC TGC TCT TCC AAG GGG TAT AAC ATT TCT TGG GAA CTA GGC AAT 734
Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu Leu Gly Asn
210 215 220

GAA CCT AAC AGT TTC CTT AAG AAG GCT GAT ATT TTC ATC AAT GGG TCG 782
Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe Ile Asn Gly Ser
225 230 235 240

CAG TTA GGA GAA GAT TAT ATT CAA TTG CAT AAA CTT CTA AGA AAG TCC 830
Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys Leu Leu Arg Lys Ser
245 250 255

ACC TTC AAA AAT GCA AAA CTC TAT GGT CCT GAT GTT GGT CAG CCT CGA 878
Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro Asp Val Gly Gln Pro Arg
260 265 270

AGA AAG ACG GCT AAG ATG CTG AAG AGC TTC CTG AAG GCT GGT GGA GAA 926
Arg Lys Thr Ala Lys Met Leu Lys Ser Phe Leu Lys Ala Gly Gly Glu
275 280 285

GTG ATT GAT TCA GTT ACA TGG CAT CAC TAC TAT TTG AAT GGA CGG ACT 974
Val Ile Asp Ser Val Thr Trp His His Tyr Tyr Leu Asn Gly Arg Thr
290 295 300

GCT ACC AGG GAA GAT TTT CTA AAC CCT GAT GTA TTG GAC ATT TTT ATT 1019
Ala Thr Arg Glu Asp Phe Leu Asn Pro Asp Val Leu Asp Ile Phe Ile
305 310 315 320

TCA TCT GTG CAA AAA GTT TTC CAG GTG GTT GAG AGC ACC AGG CCT GGC 1067
Ser Ser Val Gln Lys Val Phe Gln Val Val Glu Ser Thr Arg Pro Gly
325 330 335

AAG AAG GTC TGG TTA GGA GAA ACA AGC TCT GCA TAT GGA GGC GGA GCG 1115
Lys Lys Val Trp Leu Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala
340 345 350

CCC TTG CTA TCC GAC ACC TTT GCA GCT GGC TTT ATG TGG CTG GAT AAA 1163
Pro Leu Leu Ser Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys
355 360 365

TTG GGC CTG TCA GCC CGA ATG GGA ATA GAA GTG GTG ATG AGG CAA GTA 1211
Leu Gly Leu Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val
370 375 380

TTC TTT GGA GCA GGA AAC TAC CAT TTA GTG GAT GAA AAC TTC GAT CCT 1259
Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro
385 390 395 400

TTA CCT GAT TAT TGG CTA TCT CTT CTG TTC AAG AAA TTG GTG GGC ACC 1307
Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr
405 410 415

AAG GTG TTA ATG GCA AGC GTG CAA GGT TCA AAG AGA AGG AAG CTT CGA 1355
Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu Arg

```
                                                  VI
          420                     425                     430

GTA TAC CTT CAT TGC ACA AAC ACT GAC AAT CCA AGG TAT AAA GAA GGA 1403
Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys Glu Gly
        435                     440                     445

GAT TTA ACT CTG TAT GCC ATA AAC CTC CAT AAC GTC ACC AAG TAC TTG 1451
Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr Lys Tyr Leu
    450                     455                     460

CGG TTA CCC TAT CCT TTT TCT AAC AAG CAA GTG GAT AAA TAC CTT CTA 1499
Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp Lys Tyr Leu Leu
465                     470                     475                     480

AGA CCT TTG GGA CCT CAT GGA TTA CTT TCC AAA TCT GTC CAA CTC AAT 1547
Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys Ser Val Gln Leu Asn
                485                     490                     495

GGT CTA ACT CTA AAG ATG GTG GAT GAT CAA ACC TTG CCA CCT TTA ATG 1595
Gly Leu Thr Leu Lys Met Val Asp Asp Gln Thr Leu Pro Pro Leu Met
            500                     505                     510

GAA AAA CCT CTC CGG CCA GGA AGT TCA CTG GGC TTG CCA GCT TTC TCA 1643
Glu Lys Pro Leu Arg Pro Gly Ser Ser Leu Gly Leu Pro Ala Phe Ser
        515                     520                     525

TAT AGT TTT TTT GTG ATA AGA AAT GCC AAA GTT GCT GCT TGC ATC TGA 1691
Tyr Ser Phe Phe Val Ile Arg Asn Ala Lys Val Ala Ala Cys Ile
    530                     535                     540         543

AAA TAA AAT ATA CTA GTC CTG ACA CTG                             1718
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 824
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12

```
CTGGCAAGAA GGTCTGGTTG GGAGAGACGA GCTCAGCTTA CGGTGGCGGT GCACCCTTGC  60
TGTCCAACAC CTTTGCAGCT GGCTTTATGT GGCTGGATAA ATTGGGCCTG TCAGCCCAGA 120
TGGGCATAGA AGTCGTGATG AGGCAGGTGT TCTTCGGAGC AGGCAACTAC CACTTAGTGG 180
ATGAAAACTT TGAGCCTTTA CCTGATTACT GGCTCTCTCT TCTGTTCAAG AAACTGGTAG 240
GTCCCAGGGT GTTACTGTCA AGAGTGAAAG GCCCAGACAG GAGCAAACTC CGAGTGTATC 300
TCCACTGCAC TAACGTCTAT CACCCACGAT ATCAGGAAGG AGATCTAACT CTGTATGTCC 360
TGAACCTCCA TAATGTCACC AAGCACTTGA AGGTACCGCC TCCGTTGTTC AGGAAACCAG 420
TGGATACGTA CCTTCTGAAG CCTTCGGGGC CGGATGGATT ACTTTCCAAA TCTGTCCAAC 480
TGAACGGTCA AATTCTGAAG ATGGTGGATG AGCAGACCCT GCCAGCTTTG ACAGAAAAAC 540
CTCTCCCCGC AGGAAGTGCA CTAAGCCTGC CTGCCTTTTC CTATGGTTTT TTTGTCATAA 600
GAAATGCCAA AATCGCTGCT TGTATATGAA AATAAAAGGC ATACGGTACC CCTGAGACAA 660
AAGCCGAGGG GGGTGTTATT CATAAAACAA AACCCTAGTT TAGGAGGCCA CCTCCTTGCC 720
GAGTTCCAGA GCTTCGGGAG GGTGGGGTAC ACTTCAGTAT TACATTCAGT GTGGTGTTCT 780
CTCTAAGAAG AATACTGCAG GTGGTGACAG TTAATAGCAC TGTG                  824
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1899
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13

```
GGGAAAGCGA GCAAGGAAGT AGGAGAGAGC CGGGCAGGCG GGGCGGGGTT GGATTGGGAG  60
CAGTGGGAGG GATGCAGAAG AGGAGTGGGA GGGATGGAGG GCGCAGTGGG AGGGGTGAGG 120
AGGCGTAACG GGGCGGAGGA AAGGAGAAAA GGGCGCTGGG GCTCGGCGGG AGGAAGTGCT 180
AGAGCTCTCG ACTCTCCGCT GCGCGGCAGC TGGCGGGGGG AGCAGCCAGG TGAGCCCAAG 240
ATGCTGCTGC GCTCGAAGCC TGCGCTGCCG CCGCCGCTGA TGCTGCTGCT CCTGGGGCCG 300
CTGGGTCCCC TCTCCCCTGG CGCCCTGCCC CGACCTGCGC AAGCACAGGA CGTCGTGGAC 360
CTGGACTTCT TCACCCAGGA GCCGCTGCAC CTGGTGAGCC CCTCGTTCCT GTCCGTCACC 420
ATTGACGCCA ACCTGGCCAC GGACCCGCGG TTCCTCATCC TCCTGGGTTC TCCAAAGCTT 480
CGTACCTTGG CCAGAGGCTT GTCTCCTGCG TACCTGAGGT TTGGTGGCAC CAAGACAGAC 540
TTCCTAATTT TCGATCCCAA GAAGGAATCA ACCTTTGAAG AGAGAAGTTA CTGGCAATCT 600
CAAGTCAACC AGGATATTTG CAAATATGGA TCCATCCCTC CTGATGTGGA GGAGAAGTTA 660
CGGTTGGAAT GGCCCTACCA GGAGCAATTG CTACTCCGAG AACACTACCA GAAAAAGTTC 720
```

VII

AAGAACAGCA CCTACTCAAG AAGCTCTGTA GATGTGCTAT ACACTTTTGC AAACTGCTCA 780
GGACTGGACT TGATCTTTGG CCTAAATGCG TTATTAAGAA CAGCAGATTT GCAGTGGAAC 840
AGTTCTAATG CTCAGTTGCT CCTGGACTAC TGCTCTTCCA AGGGGTATAA CATTTCTTGG 900
GAACTAGGCA ATGAACCTAA CAGTTTCCTT AAGAAGGCTG ATATTTTCAT CAATGGGTCG 960
CAGTTAGGAG AAGATTATAT TCAATTGCAT AAACTTCTAA GAAAGTCCAC CTTCAAAAAT 1020
GCAAAACTCT ATGGTCCTGA TGTTGGTCAG CCTCGAAGAA AGACGGCTAA GATGCTGAAG 1080
AGCTTCCTGA AGGCTGGTGG AGAAGTGATT GATTCAGTTA CATGGCATCA CTACTATTTG 1140
AATGGACGGA CTGCTACCAG GGAAGATTTT CTAAACCCTG ATGTATTGGA CATTTTTATT 1200
TCATCTGTGC AAAAAGTTTT CCAGGTGGTT GAGAGCACCA GGCCTGGCAA GAAGGTCTGG 1260
TTAGGAGAAA CAAGCTCTGC ATATGGAGGC GGAGCGCCCT TGCTATCCGA CACCTTTGCA 1320
GCTGGCTTTA TGTGGCTGGA TAAATTGGGC CTGTCAGCCC GAATGGGAAT AGAAGTGGTG 1380
ATGAGGCAAG TATTCTTTGG AGCAGGAAAC TACCATTTAG TGGATGAAAA CTTCGATCCT 1440
TTACCTGATT ATTGGCTATC TCTTCTGTTC AAGAAATTGG TGGGCACCAA GGTGTTAATG 1500
GCAAGCGTGC AAGGTTCAAA GAGAAGGAAG CTTCGAGTAT ACCTTCATTG CACAAACACT 1560
GACAATCCAA GGTATAAAGA AGGAGATTTA ACTCTGTATG CCATAAACCT CCATAACGTC 1620
ACCAAGTACT TGCGGTTACC CTATCCTTTT TCTAACAAGC AAGTGGATAA ATACCTTCTA 1680
AGACCTTTGG GACCTCATGG ATTACTTTCC AAATCTGTCC AACTCAATGG TCTAACTCTA 1740
AAGATGGTGG ATGATCAAAC CTTGCCACCT TTAATGGAAA AACCTCTCCG GCCAGGAAGT 1800
TCACTGGGCT TGCCAGCTTT CTCATATAGT TTTTTTGTGA TAAGAAATGC CAAAGTTGCT 1860
GCTTGCATCT GAAAATAAAA TATACTAGTC CTGACACTG 1899

(2) INFORMATION FOR SEQ ID NO:14:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 592
(B) TYPE: amino acid
(C) STRANDEDNESS: singl
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14

Met Glu Gly Ala Val Gly Gly Val Arg Arg Arg Asn Gly Ala Glu
5 10 15
Glu Arg Arg Lys Gly Arg Trp Gly Ser Ala Gly Gly Ser Ala Arg
20 25 30
Ala Leu Asp Ser Pro Leu Arg Gly Ser Trp Arg Gly Glu Gln Pro
35 40 45
Gly Glu Pro Lys Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro
50 55 60
Pro Leu Met Leu Leu Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro
65 70 75
Gly Ala Leu Pro Arg Pro Ala Gln Ala Gln Asp Val Val Asp Leu
80 85 90
Asp Phe Phe Thr Gln Glu Pro Leu His Leu Val Ser Pro Ser Phe
95 100 105
Leu Ser Val Thr Ile Asp Ala Asn Leu Ala Thr Asp Pro Arg Phe
110 115 120
Leu Ile Leu Leu Gly Ser Pro Lys Leu Arg Thr Leu Ala Arg Gly
125 130 135
Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly Thr Lys Thr Asp Phe
140 145 150
Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe Glu Glu Arg Ser
155 160 165
Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys Tyr Gly Ser
170 175 180
Ile Pro Pro Asp Val Glu Glu Lys Leu Arg Leu Glu Trp Pro Tyr
185 190 195
Gln Glu Gln Leu Leu Leu Arg Glu His Tyr Gln Lys Lys Phe Lys
200 205 210
Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe
215 220 225
Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu
230 235 240
Leu Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu
245 250 255
Leu Leu Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu
260 265 270
Leu Gly Asn Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe
275 280 285
Ile Asn Gly Ser Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys
290 295 300
Leu Leu Arg Lys Ser Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro
305 310 315
Asp Val Gly Gln Pro Arg Arg Lys Thr Ala Lys Met Leu Lys Ser
320 325 330
Phe Leu Lys Ala Gly Gly Glu Val Ile Asp Ser Val Thr Trp His
335 340 345

VIII

His Tyr Tyr Leu Asn Gly Arg Thr Ala Thr Arg Glu Asp Phe Leu
350 355 360
Asn Pro Asp Val Leu Asp Ile Phe Ile Ser Ser Val Gln Lys Val
365 370 375
Phe Gln Val Val Glu Ser Thr Arg Pro Gly Lys Lys Val Trp Leu
380 385 390
Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala Pro Leu Leu Ser
395 400 405
Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys Leu Gly Leu
410 415 420
Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val Phe Phe
425 430 435
Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro Leu
440 445 450
Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr
455 460 465
Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu
470 475 480
Arg Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys
485 490 495
Glu Gly Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr
500 505 510
Lys Tyr Leu Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp
515 520 525
Lys Tyr Leu Leu Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys
530 535 540
Ser Val Gln Leu Asn Gly Leu Thr Leu Lys Met Val Asp Asp Gln
545 550 555
Thr Leu Pro Pro Leu Met Glu Lys Pro Leu Arg Pro Gly Ser Ser
560 565 570
Leu Gly Leu Pro Ala Phe Ser Tyr Ser Phe Phe Val Ile Arg Asn
575 580 585
Ala Lys Val Ala Ala Cys Ile
590 592

(2) INFORMATION FOR SEQ ID NO:15:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1899
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15

GGG 3
AAA GCG AGC AAG GAA GTA GGA GAG AGC CGG GCA GGC GGG GCG GGG 48
TTG GAT TGG GAG CAG TGG GAG GGA TGC AGA AGA GGA GTG GGA GGG 93
ATG GAG GGC GCA GTG GGA GGG GTG AGG AGG CGT AAC GGG GCG GAG 138
Met Glu Gly Ala Val Gly Gly Val Arg Arg Arg Asn Gly Ala Glu
5 10 15

GAA AGG AGA AAA GGG CGC TGG GGC TCG GCG GGA GGA AGT GCT AGA 183
Glu Arg Arg Lys Gly Arg Trp Gly Ser Ala Gly Gly Ser Ala Arg
20 25 30

GCT CTC GAC TCT CCG CTG CGC GGC AGC TGG CGG GGG GAG CAG CCA 228
Ala Leu Asp Ser Pro Leu Arg Gly Ser Trp Arg Gly Glu Gln Pro
35 40 45

GGT GAG CCC AAG ATG CTG CTG CGC TCG AAG CCT GCG CTG CCG CCG 273
Gly Glu Pro Lys Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro
50 55 60

CCG CTG ATG CTG CTG CTC CTG GGG CCG CTG GGT CCC CTc TCC CCT 318
Pro Leu Met Leu Leu Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro
65 70 75

GGC GCC CTG CCC CGA CCT GCG CAA GCA CAG GAC GTc GTG GAC CTG 363
Gly Ala Leu Pro Arg Pro Ala Gln Ala Gln Asp Val Val Asp Leu
80 85 90

GAC TTc TTC ACC CAG GAG CCG CTG CAC CTG GTG AGC CCC TCG TTC 408
Asp Phe Phe Thr Gln Glu Pro Leu His Leu Val Ser Pro Ser Phe
95 100 105

CTG TCC GTC ACC ATT GAC GCC AAC CTG GCC ACG GAC CCG CGG TTC 453

IX
Leu Ser Val Thr Ile Asp Ala Asn Leu Ala Thr Asp Pro Arg Phe
110 115 120

CTC ATC CTC CTG GGT TCT CCA AAG CTT CGT ACC TTG GCC AGA GGC 498
Leu Ile Leu Leu Gly Ser Pro Lys Leu Arg Thr Leu Ala Arg Gly
125 130 135

TTG TCT CCT GCG TAC CTG AGG TTT GGT GGC ACC AAG ACA GAC TTC 543
Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly Thr Lys Thr Asp Phe
140 145 150

CTA ATT TTC GAT CCC AAG AAG GAA TCA ACC TTT GAA GAG AGA AGT 588
Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe Glu Glu Arg Ser
155 160 165

TAC TGG CAA TCT CAA GTC AAC CAG GAT ATT TGC AAA TAT GGA TCC 633
Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys Tyr Gly Ser
170 175 180

ATC CCT CCT GAT GTG GAG GAG AAG TTA CGG TTG GAA TGG CCC TAC 678
Ile Pro Pro Asp Val Glu Glu Lys Leu Arg Leu Glu Trp Pro Tyr
185 190 195

CAG GAG CAA TTG CTA CTC CGA GAA CAC TAC CAG AAA AAG TTC AAG 723
Gln Glu Gln Leu Leu Leu Arg Glu His Tyr Gln Lys Lys Phe Lys
200 205 210

AAC AGC ACC TAC TCA AGA AGC TCT GTA GAT GTG CTA TAC ACT TTT 768
Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe
215 220 225

GCA AAC TGC TCA GGA CTG GAC TTG ATC TTT GGC CTA AAT GCG TTA 813
Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu
230 235 240

TTA AGA ACA GCA GAT TTG CAG TGG AAC AGT TCT AAT GCT CAG TTG 858
Leu Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu
245 250 255

CTC CTG GAC TAC TGC TCT TCC AAG GGG TAT AAC ATT TCT TGG GAA 903
Leu Leu Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu
260 265 270

CTA GGC AAT GAA CCT AAC AGT TTC CTT AAG AAG GCT GAT ATT TTC 948
Leu Gly Asn Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe
275 280 285

ATC AAT GGG TCG CAG TTA GGA GAA GAT TAT ATT CAA TTG CAT AAA 993
Ile Asn Gly Ser Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys
290 295 300

CTT CTA AGA AAG TCC ACC TTC AAA AAT GCA AAA CTC TAT GGT CCT 1038
Leu Leu Arg Lys Ser Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro
305 310 315

GAT GTT GGT CAG CCT CGA AGA AAG ACG GCT AAG ATG CTG AAG AGC 1083
Asp Val Gly Gln Pro Arg Arg Lys Thr Ala Lys Met Leu Lys Ser
320 325 330

TTC CTG AAG GCT GGT GGA GAA GTG ATT GAT TCA GTT ACA TGG CAT 1128
Phe Leu Lys Ala Gly Gly Glu Val Ile Asp Ser Val Thr Trp His
335 340 345

CAC TAC TAT TTG AAT GGA CGG ACT GCT ACC AGG GAA GAT TTT CTA 1173
His Tyr Tyr Leu Asn Gly Arg Thr Ala Thr Arg Glu Asp Phe Leu
350 355 360

AAC CCT GAT GTA TTG GAC ATT TTT ATT TCA TCT GTG CAA AAA GTT 1218
Asn Pro Asp Val Leu Asp Ile Phe Ile Ser Ser Val Gln Lys Val
365 370 375

TTC CAG GTG GTT GAG AGC ACC AGG CCT GGC AAG AAG GTC TGG TTA 1263
Phe Gln Val Val Glu Ser Thr Arg Pro Gly Lys Lys Val Trp Leu
380 385 390

X
GGA GAA ACA AGC TCT GCA TAT GGA GGC GGA GCG CCC TTG CTA TCC 1308
Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala Pro Leu Leu Ser
395 400 405

GAC ACC TTT GCA GCT GGC TTT ATG TGG CTG GAT AAA TTG GGC CTG 1353
Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys Leu Gly Leu
410 415 420

TCA GCC CGA ATG GGA ATA gAA GTG GTG ATG AGG CAA GTA TTC TTT 1398
Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val Phe Phe
425 430 435

GGA GCA GGA AAC TAC CAT TTA GTG GAT GAA AAC TTC GAT CCT TTA 1443
Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro Leu
440 445 450

CCT GAT TAT TGG CTA TCT CTT CTG TTC AAG AAA TTG GTG GGC ACC 1488
Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr
455 460 465

AAG GTG TTA ATG GCA AGC GTG CAA GGT TCA AAG AGA AGG AAG CTT 1533
Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu
470 475 480

CGA GTA TAC CTT CAT TGC ACA AAC ACT GAC AAT CCA AGG TAT AAA 1578
Arg Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys
485 490 495

GAA GGA GAT TTA ACT CTG TAT GCC ATA AAC CTC CAT AAC GTC ACC 1623
Glu Gly Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr
500 505 510

AAG TAC TTG CGG TTA CCC TAT CCT TTT TCT AAC AAG CAA GTG GAT 1668
Lys Tyr Leu Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp
515 520 525

AAA TAC CTT CTA AGA CCT TTG GGA CCT CAT GGA TTA CTT TCC AAA 1713
Lys Tyr Leu Leu Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys
530 535 540

TCT GTC CAA CTC AAT GGT CTA ACT CTA AAG ATG GTG GAT GAT CAA 1758
Ser Val Gln Leu Asn Gly Leu Thr Leu Lys Met Val Asp Asp Gln
545 550 555

ACC TTG CCA CCT TTA ATG GAA AAA CCT CTC CGG CCA GGA AGT TCA 1803
Thr Leu Pro Pro Leu Met Glu Lys Pro Leu Arg Pro Gly Ser Ser
560 565 570

CTG GGC TTG CCA GCT TTC TCA TAT AGT TTT TTT GTG ATA AGA AAT 1848
Leu Gly Leu Pro Ala Phe Ser Tyr Ser Phe Phe Val Ile Arg Asn
575 580 585

GCC AAA GTT GCT GCT TGC ATC TGA AAA TAA AAT ATA CTA GTC CTG 1893
Ala Lys Val Ala Ala Cys Ile
590 592

ACA CTG 1899

(2) INFORMATION FOR SEQ ID NO:16:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 594
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16

ATTACTATAG GGCACGCGTG GTCGACGGCC CGGGCTGGTA TTGTCTTAAT GAGAAGTTGA 60
TAAAGAATTT TGGGTGGTTG ATCTCTTTCC AGCTGCAGTT TAGCGTATGC TGAGGCCAGA 120
TTTTTTCAGG CAAAAGTAAA ATACCTGAGA AACTGCCTGG CCAGAGGACA ATCAGATTTT 180
GGCTGGCTCA AGTGACAAGC AAGTGTTTAT AAGCTAGATG GGAGAGGAAG GGATGAATAC 240
TCCATTGGAG GCTTTACTCG AGGGTCAGAG GGATACCCGG CGCCATCAGA ATGGGATCTG 300
GGAGTCGGAA ACGCTGGGTT CCCACGAGAG CGCGCAGAAC ACGTGCGTCA GGAAGCCTGG 360
TCCGGGATGC CCAGCGCTGC TCCCCGGGCG CTCCTCCCCG GGCGCTCCTC CCCAGGCCTC 420
CCGGGCGCTT GGATCCCGGC CATCTCCGCA CCCTTCAAGT GGGTGTGGGT GATTTCGTAA 480
GTGAACGTGA CCGCCACCGG GGGGAAAGCG AGCAAGGAAG TAGGAGAGAG CCGGGCAGGC 540
GGGGCGGGGT TGGATTGGGA GCAGTGGGAG GGATGCAGAA GAGGAGTGGG AGGG 594

(2) INFORMATION FOR SEQ ID NO:17:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17
CCCCAGGAGC AGCAGCATCA G 21

(2) INFORMATION FOR SEQ ID NO:18:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18
AGGCTTCGAG CGCAGCAGCA T 21

(2) INFORMATION FOR SEQ ID NO:19:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19
GTAATACGAC TCACTATAGG GC 22

(2) INFORMATION FOR SEQ ID NO:20:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20
ACTATAGGGC ACGCGTGGT 19

(2) INFORMATION FOR SEQ ID NO:21:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21
CTTGGGCTCA CCTGGCTGCT C 21

(2) INFORMATION FOR SEQ ID NO:22:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22
AGCTCTGTAG ATGTGCTATA CAC 23

(2) INFORMATION FOR SEQ ID NO:23:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23
GCATCTTAGC CGTCTTTCTT CG 22

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/17954

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12N 15/56, 15/63, 1/21, 9/24, 15/11; A61K 38/47
US CL :536/23.1, 23.2; 435/200, 325, 252.3, 320.1; 424/94.61
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.1, 23.2; 435/200, 325, 252.3, 320.1; 424/94.61

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, SCISEARCH, BIOSIS, EMBASE, WPI, BIOTECHDS, NTIS, CA, LIFESCI
search terms: heparanase#, gene# or sequence#

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X - Y	US 5,362,641 A (FUKS et al.) 08 November 1994, see entire document	28, 29, 33-35, 37,38 -------- 1,8,9,11,18,19,26,27,36,39-41
X	WO 95/04158 A1 (UPJOHN CO.) 09 February 1995, see entire document.	1, 8, 11, 18, 19, 26-29, 33, 34-38
X	Database GenBank on STN, US National Library of Medicine (Bethesda MD), HILLIER et al., 'The WashU-Merck EST Project, No. N32056, 10 January 1996.	9, 10

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 NOVEMBER 1998	11 JAN 1999

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	D. Lawrence Foe REBECCA PROUTY
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/17954

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database GenBank on STN, US National Library of Medicine (Bethesda MD), No. 30845, HILLIER et al., 'The WashU-Merck EST Project, 05 January 1996	9, 10
X	Database GenBank on STN, US National Library of Medicine (Bethesda MD), HILLIER et al., 'The WashU-Merck EST Project. No. N30824, 05 January 1996.	9, 10
X	Database GenBank on STN, National Library of Medicine (Bethesda MD), ADAMS et al., ' Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. No. AA304653, 18 April 1997.	30